Essays on Adaptive Learning with Applications to Monetary Policy

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Oliver Fries
aus Saarlouis.

Tag der Promotion: 6. Mai 2008

Dekan:	Prof. Dr. Erik Theissen
Erstreferent:	Prof. Dr. Jürgen von Hagen
Zweitreferent:	Prof. Dr. Ludger Linnemann

Tag der mündlichen Prüfung: 6. Mai 2008

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Acknowledgements

This thesis presents the main results of my research at the Bonn Graduate School of Economics (BGSE) and the Center for European Integration Studies (ZEI). I am indebted to a number of people who supported me and my work during the past four years. Firstly, I would like to thank my supervisor Prof. Dr. Jürgen von Hagen for his guidance and support. I benefited much from his comments and the stimulating research environment he provided at the ZEI. Especially important for my studies were the yearly ZEI summerschools on Monetary Theory and Policy. I also want to thank my colleagues Simon Junker, Michael Evers and Michael Massmann for numerous fruitful discussions on various economic and econometric topics. I would like to acknowledge the help of Prof. Dr. Bernard Delyon regarding the application of his stochastic approximation results. I am grateful to all colleagues from the University of Bonn, the participants of the BGSE workshops and the ZEI summerschools for various comments and suggestions. Moreover, I want to thank Prof. Dr. Urs Schweizer, Prof. Dr. Georg Nöldecke and again my supervisor Prof. Dr. Jürgen von Hagen for their combined efforts in running the BGSE. The material support of the BGSE, the DFG and the ZEI is also gratefully acknowledged. Finally I want to especially thank Santina, my parents and the rest of my family for their strong support and encouragement.

Oliver Fries
October 2007, Bonn

Contents

List of Figures

Introduction

While the concept of rational expectations has become the standard tool of modelling expectations in macroeconomic models, it has been criticised for its high information and rationality requirements. Moreover, rational expectations is an equilibrium concept which does not answer the question how the equilibrium can be reached. As a consequence there has been an increasing interest in learning theory in economics in recent years. This thesis focuses on a particular concept of learning which is called adaptive learning.

Rational expectations are often criticised to require too much information on part of the economic agents. It is assumed that they know the structure of the economy, the history of all endogenous and exogenous variables and all parameter values. The literature on adaptive learning reduces these informational requirements. The agents are still assumed to know the history of all variables. It is also assumed that they are endowed with some model of the economy which they use to forecast the future path of the endogenous variables. This forecasting model can be a an arbitrary function of past endogenous and past and current exogenous variables. In the literature it is usually referred to as the perceived law of motion of the economy. Unlike under rational expectations the agents do not know the correct values of their perceived law of motion. Instead they try to infer these parameters from the data they observe using some learning algorithm. Popular adaptive learning algorithms include recursive least squares and stochastic gradient learning. Given some estimates for the perceived law of motion the agents use the perceived law of motion to forecast the future path of the endogenous variables. These expectations together with the history of endogenous and exogenous variables determine the actual law of motion of the economy. Subsequently a new set of exogenous variables realises. The new realisation of endogenous and exogenous variables enables the agents to update the estimates for their perceived law of motion and the process starts all over. Under certain conditions on the structure of the economy and the perceived law of motion the forecasts of the agents converge towards rational expectations.

This thesis contributes to three related branches of the literature, which are described in the following. The first investigates the conditions under which convergence occurs. Once we know these conditions we can investigate whether the design of monetary or fiscal policy has an impact on the convergence of adaptive learning. Finally, adaptive learning may be used to explain empirical observations which are difficult to reconcile with rational expectations.

I.1. Convergence of Adaptive Learning and E-stability

One of the first papers on least squares learning was the paper by Bray and Savin (1986), who investigate the convergence of least squares learning towards rational expectations in the cobweb model. The paper demonstrates that even in the very simple cobweb model it is extremely difficult to prove the convergence of an adaptive learning algorithm. Initiated by the work of Marcet and Sargent (1989a,b) it has been shown in a sequence of papers that convergence of least squares learning towards rational expectations is equivalent to the stability of a certain differential equation (ODE), which is called the mean dynamics or mean ODE.[1]

A collection of different convergence results and various applications can be found in Evans and Honkapohja (2001). All these convergence results relate the problem of convergence of an adaptive learning algorithm like recursive least squares to the local stability of the mean ODE. It is often possible to prove that the mean ODE is stable if and only if a smaller and simpler ODE is stable. The stability of the smaller ODE has become known as E-stability. In most cases it is possible to apply standard stability results for differential equations to investigate E-stability. The conditions which have to be checked and the computational effort are then comparable to the analysis of local determinacy. Chapter 1 presents an application of these conditions to a model of monetary and fiscal policy.

Under particular circumstances it can however be very hard to check E-stability. Technically this is the case whenever the ODE defining E-stability is non-hyperbolic. Economically this is the case if the equilibrium is not locally unique, for example due to the presence of sunspots. This is particulary disturbing because it is controversial in the economic literature whether sunspots are worrisome or whether they are simply a mathematical artefact. Therefore it would be very desirable to know whether sunspot equilibria are learnable or not. Apart from the problems with the analysis of E-stability there is no prove in the literature so far which shows that E-stability governs the convergence of least squares learning also if the equilibrium is not locally unique. Chapter 2 presents new results on E-stability and the convergence of least squares learning for two prominent representations of locally non-unique sunspot equilibria, general form and common factor representations. Our results support the hypothesis which has been formulated in the literature that E-stability continues to determine the convergence of least squares learning even in the presence of locally non-unique sunspot equilibria.

[1] The equivalence is subject to certain regularity conditions like stationarity of the equilibrium.

Under least squares learning the weight on a new observation decreases with the total number of observations. This is because ordinary least squares assumes that the coefficients which are estimated remain constant over time. If the agents doubt that the structure is indeed constant, then they may want to put more weight on recent observations to track structural changes. In most cases these structural changes are breaks in the conduct of monetary or fiscal policy. Chapter 3 shows that least squares learning with a constant gain is able to track breaks in monetary policy. A theoretical justification for least squares with a constant gain has been brought forward by Sargent and Williams (2005). They show that recursive least squares with a constant gain is approximately optimal if the coefficients to be estimated follow a random walk.

Another interesting phenomenon may arise if some agents, e.g. the households or the central bank, have a misspecified perceived law of motion. In these cases the economy may still converge to a so called self-confirming equilibrium. In a self-confirming equilibrium the new data which becomes available to the agents validates their beliefs, so that they cannot detect the misspecification given their model of the economy. If in addition the agents use a constant gain learning algorithm, then escapes may occur. That is there may be large and recurring departures from the self-confirming equilibrium. Given the misspecification and the fact that agents try to track structural changes it is perhaps not too surprising that escapes do occur. The remarkable thing about escapes is that the escape path is predictable. An escape can be triggered by a sequence of large shocks which drive the beliefs away from the self-confirming equilibrium in the direction of the most likely escape path. In order to achieve this the shocks must be correlated in the right way. A detailed description of the technical details can be found in Williams (2001).

I.2. E-Stability and Policy Design

A large part of the literature on adaptive learning investigates the E-stability of rational expectations equilibria across different economic models. It is often argued that any reasonable equilibrium should be learnable by adaptively learning agents. Equilibria which are E-unstable cannot be attained by adaptively learning agents and so they can be completely disregarded. This argument uses adaptive learning as a tool for equilibrium justification. However, in many models the equilibrium depends on parameters chosen by some policymaker and often the policymaker can implement an E-stable equilibrium by choosing an appropriate policy. The policymaker should then choose an E-stable equilibrium to avoid unnecessary fluctuations in the expectations of the agents which lead to higher volatility in aggregate economic variables like output and inflation.

E-stability is a local property of a model and for most applications it is not possible to derive global stability results. In order to exclude the possibility that there are multiple learnable equilibria it is useful to check both E-stability and determinacy simultaneously. Perhaps the most influential paper which analyses E-stability and determinacy jointly is Bullard and Mitra (2002), who investigate the effect of different interest rate rules in a New Keynesian model. Their main result is that monetary policy should follow the Taylor principle in order to implement a determinate and E-stable equilibrium. In Bullard and Mitra (2007) they demonstrate that inertia can help alleviate problems of indeterminacy and promote E-stability.

Chapter 1 compares the E-stability and determinacy properties of monetarist and fiscalist equilibria. There is a dispute in the literature about the plausibility of fiscalist equilibria. Many authors doubt the feasibility of fiscalist equilibria, but only recently McCallum (2001) claimed that they are implausible because they are not attainable by adaptively learning agents. Evans and Honkapohja (2004) however present a flexible price model where both fiscalist and monetarist equilibria may be stable under learning. We analyse this issue in chapter 1 in a standard New Keynesian model. Our results support the view that fiscalist equilibria are indeed stable under learning for certain combinations of monetary and fiscal policy. We also show that a failure of policy coordination may lead to instability under learning.

Another important result of the literature on E-stability and determinacy is the instability of fundamentals based interest rate rules. As shown by Evans and Honkapohja (2003) all interest rate rules which condition only on exogenous variables lead to E-instability and indeterminacy. Especially this is true for the optimal policy under discretion. It is however possible to write the optimal rule under discretion as a function not only of the exogenous variables but also of expected inflation and the expected output gap. It turns out that this alternative formulation leads to both determinacy and E-stability (cf. Evans and Honkapohja, 2003).

I.3. Empirical Applications of Adaptive Learning

Recently there has been a growing interest in the additional dynamics introduced by learning. The majority of this part of the learning literature assumes that the agents in the economy use some constant gain learning algorithm, which is able to track structural changes and leads to perpetual learning. As noted above recursive least squares places a decreasing weight on new observations. We have also discussed that least squares learning has been shown to converge to an E-stable rational expectations equilibrium for a large class of models. A prerequisite for this is however that the structure of the economy remains constant over time. In contrast a constant gain algorithm puts more weight on recent information and an exponentially declining weight on old information. Constant gain algorithms can be justified as an approximately optimal algorithm provided the coefficients to be estimated follow a random walk. In the absence of any breaks it can be shown (again under some regularity assumptions on the model and the learning algorithm) that constant gain algorithms converge in distribution to an E-stable equilibrium. For the agents the advantage of those algorithms is that they can track parameter changes, while the cost to pay for that is that the algorithm converges only in distribution to an E-stable equilibrium, which means that the fluctuations around the equilibrium will never disappear. From the point of view of an economist constant gain algorithms lead to perpetual learning. Thus constant gain learning may explain repeated and substantial deviations from rational expectations as well as the transition from one equilibrium to another subsequent to a structural break.

Some important contributions in this area of research are Cho et al. (2002) and Primiceri (2005) who model the Great Inflation in the US during the 70s and 80s using escape dynamics. A crucial difference between both papers is that in Cho et al. (2002) the escape is from high to low inflation, so as time goes by inflation should return to the high levels of the 70s and 80s. In contrast Primiceri (2005) presents a model where the high levels of inflation are the escapes, so that the low level of inflation we experience at the moment represents the equilibrium. Bullard and Cho (2005) build a model of the liquidity trap in Japan also using escape dynamics. Adam et al. (2006) and Guidolin and Timmermann (2007) investigate whether adaptive learning can explain stylised facts from the asset pricing literature which are hard to reconcile with the standard models under rational expectations. Their results indicate that adaptive learning improves the ability to replicate these stylised facts but the extent to which this is case is still controversial. Chapter 3 demonstrates that the Great Moderation, that is the long and large decline of aggregate economic volatility over the last decades, can be modelled as the transition under learning from an equilibrium with high volatility to low volatility in response to a break in monetary policy under Fed chairman Paul Volcker.

I.4. Outlook

The rest of this thesis is divided in three chapters each of which is self-contained. Chapter 1 studies the determinacy and learnability of rational expectations equilibria in a New Keynesian model with interacting authorities. Monetary policy controls the nominal interest rate via a linear feedback rule that depends on lagged, contemporaneous or forward data of inflation and the output gap. Taxes are set as a linear function of real government debt. There are two kinds of equilibria, monetarist and fiscalist equilibria. We find that learnability requires that both authorities coordinate their policies. If monetary policy is active and fiscal policy is passive the orthodox monetarist equilibrium is the unique learnable equilibrium. If on the other hand fiscal policy is active and monetary policy is passive then there is a unique learnable fiscalist equilibrium. In case of lagged or forward data in the interest rate rule determinacy and learnability require in addition that the nominal interest rate is not too sensitive to fluctuations in the output gap. These results demonstrate that in contrast to the claim of McCallum (2001) adaptive learning may converge to a fiscalist equilibrium in one of the most widely used models for monetary policy analysis. Chapter 1 contributes to the literature on E-stability and determinacy. It applies E-stability both as a justification for fiscalist equilibria as well as to give policy recommendations in an environment where both monetary and fiscal policies are rule based.

Chapter 2 investigates the stability under learning of general form and common factor sunspot representations in self-referential multivariate linear models. It establishes necessary and sufficient conditions for E-stability of common factor sunspot representations and proves an instability result for general form representations. It is shown that the convergence of least squares learning to a sunspot representation (augmented with projection facilities) is governed by the stability of the mean ODE. In a standard New Keynesian model with a forward looking interest rate rule E-stable common factor sunspot representations exist if the nominal interest rate is very sensitive to changes in the output gap. In our calibrated economy least squares learning converges to an E-stable common factor representation even without imposition of a projection facility. Our results support the conjecture in the literature that under standard regularity conditions E-stability and convergence of least squares learning are equivalent not only for fundamental equilibria but also in the presence of sunspots. Thus chapter 2 extends the literature on the convergence of adaptive learning and E-stability.

Over the last decades most industrialised countries have experienced a tremendous decline in the volatilities of both inflation and aggregate economic activity. Chapter 3 builds a microfounded model of inflation and the output gap in which information is sticky. The nominal interest rate is set as a linear feedback rule of inflation and the output gap. The information acquisition rate is chosen endogenously by adaptively learning agents. We analyse the consequences of a shift in monetary policy towards more inflation stabilisation for the volatilities of inflation and the output gap. A simulation exercise shows that in the long run the model may generate a decline in both volatilities comparable to US data even if there is a significant volatility trade-off in the short run. Under adaptive learning agents change their information acquisition rate only gradually. In the short run the information acquisition rate is essentially fixed and the well-known volatility trade-off prevails. However, in the long run the shift in monetary policy leads to a reduction of the information acquisition rate. Thus information disseminates more slowly and the volatility trade-off may break down.

CHAPTER 1

Learning about Monetary and Fiscal Policy Rules

1.1. Introduction

A policy rule can be defined as 'nothing more than a systematic decision process that uses information in a consistent and predictable way (Meltzer, 1993, p.223)'. The main benefit of using a credible policy rule instead of discretion is that it allows the policymaker to influence the expectations of the agents in the economy so that inflation and output can be stabilised more efficiently. Although monetary policy rules play a prominent role in modern macroeconomics this is not the case for fiscal policy rules. Most of the literature on rule based monetary policy assumes that fiscal policy simply balances the budget. Probably the reason lies in the predominant view in the literature that inflation is always and everywhere a monetary phenomenon. However, modelling both monetary and fiscal policy as being rule based leads to the conclusion that in some circumstances fiscal variables like real government debt are equally important determinants of inflation (cf. Leeper, 1991). As a result there has been a controversial debate about the determinants of inflation in recent years.

On the one hand there are the proponents of the traditional view that inflation is a monetary phenomenon. While these monetarists recognise that a well behaved fiscal policy is important to achieve price stability, they conjecture that a strong monetary policy authority can compel fiscal policy not to misbehave. Especially fiscal policy cannot follow a policy rule which is independent of monetary policy. Monetarists advocate strong independent central banks with a focus on price stability. In their view this is enough to assure price stability as fiscal policy has to adapt to fulfill the budget constraint (cf. Sargent and Wallace, 1981).

The diametrically opposed group are the advocates of the fiscal theory of the price level (FTPL), the so called fiscalists.[1] They believe that an independent inflation fighting central bank is not enough to assure price stability. The conflict with the monetarists arises because the fiscalists don't believe that the government budget constraint has to hold for all possible values of the price level. In their view the budget constraint is an equilibrium condition, so given some monetary and fiscal policy, the price level adapts to satisfy the

[1] A comprehensive survey of the FTPL is Christiano and Fitzgerald (2000).

budget constraint. Thus both monetary and fiscal policy may follow independent policy rules and price stability hinges on the coordination of monetary and fiscal policy.

A prominent case of the FTPL states that the price level is entirely determined by fiscal variables if monetary policy fixes the money supply (cf. e.g. Sims (1999) or McCallum (2001)). This is the most unorthodox example of the FTPL. McCallum (2001) develops a model for this type of the FTPL and argues that it is implausible for two reasons. First, while the fiscalist solution exists, it is a bubble solution whereas the orthodox monetarist solution represents the MSV solution. More fundamental in his view the monetarist solution is learnable by adaptively learning agents whereas the fiscalist solution is not. [2]

The models of Sims (1999) and McCallum (2001) describe the most extreme version of the FTPL. There are less radical models of the FTPL where in equilibrium both monetary as well as fiscal variables affect inflation. In these cases it is important to distinguish between active and passive monetary and fiscal policy. According to Leeper (1991) monetary policy is active (AM) if it leans against inflation and raises interest rates more than one for one with inflation and passive (PM) otherwise. This condition is also known as the Taylor principle. Fiscal policy is said to be passive (PF) if an increase in government debt leads to an increase in taxes larger than the real interest paid on the additional debt and active (AF) otherwise.[3] Leeper (1991) showed that there are stationary equilibria where fiscal policy is active, monetary policy is passive and inflation depends on both monetary and fiscal policy. Specifically he showed that there is a determinate REE if and only if monetary policy is active and fiscal policy is passive or vice versa. The distinguishing feature of the determinate solutions under active fiscal and passive monetary policy is that inflation depends on the level of outstanding real government debt. Therefore they are called fiscalist solutions in contrast to the orthodox monetarist solutions, in which inflation is independent of fiscal variables. The distinction between fiscalist and monetarist solutions follows McCallum (2001) and Evans and Honkapohja (2004). [4]

Following Bullard and Mitra (2002) we analyse the determinacy and learnability of rational expectations equilibria in a standard new Keynesian model for different interest rate rules that respond to lagged data, contemporaneous data and forward expectations of inflation and the output gap. Bullard and Mitra (2002) follow the standard procedure and assume that fiscal policy balances the budget using lump sum taxes and is thus irrelevant for

[2] Woodford (2003a) argues that the E-instability of the fiscalist solution found by McCallum (2001) results from the fact that McCallum uses a restricted class of perceived laws of motion that exclude the possibility of a fiscalist solution by assumption.

[3] A precise definition of active and passive monetary and fiscal policy is given in definition 1.4.1 for monetary policy and in definition 1.5.1 for fiscal policy.

[4] The definition of monetarist and fiscalist solutions (cmp. definition 1.5.2) follows Evans and Honkapohja (2004).

the analysis of determinacy and learnability. In contrast, we assume that the government sets lump sum taxes using a linear feedback rule of real government debt as proposed by Leeper (1991). As a consequence fiscal policy can no longer be disregarded and will have important effects on both determinacy and learnability. Especially, it allows for the possibility of both active and passive fiscal policy and thus the existence of fiscalist equilibria.

As mentioned above the model of Leeper (1991) covers a wide range of combinations of monetary and fiscal policy. Especially there are stationary equilibria in which both monetary and fiscal variables affect inflation. So the model is a less extreme version of the FTPL than the one investigated by McCallum (2001). Following the instability result of McCallum (2001), Evans and Honkapohja (2004) examine learnability in a model that is close to the one presented by Leeper (1991) and find that the fiscalist solution is learnable under active fiscal and passive monetary policy. If both policies are active the equilibrium is explosive. Yet if monetary policy is sufficiently active the monetarist equilibrium is still learnable. Under passive fiscal and active monetary policy they find that the monetarist solution is learnable. However the analysis is based on a constant endowment economy with flexible prices. Additionally the monetary policy authority is restricted to use a monetary feedback rule that depends on current period inflation. While this is a sensible starting point and a simple model to derive analytical results, it is not clear how the results change in models with price stickiness or when the monetary policy authority uses different policy rules. It is well known that determinacy and E-stability are very sensitive to the specific type of feedback rule employed by the central bank and that output stabilisation is a crucial determinant for determinacy and learnability (cf. Bullard and Mitra, 2002). Therefore this chapter extends the work of Evans and Honkapohja (2004) to a standard New Keynesian model with an intertemporal IS-curve, where the monetary feedback rule may depend on both inflation and the output gap.

It turns out that under passive fiscal policy only monetarist equilibria may be learnable, whereas under active fiscal policy only fiscalist equilibria can be learnable. A determinate and learnable equilibrium obtains if monetary policy is active (i.e. the Taylor principle holds) and the nominal interest rate is not too sensitive to changes in the output gap. Under active fiscal the nominal interest rate still should not be too sensitive too output gap changes but in addition monetary policy needs to be passive to assure that there is a determinate and learnable equilibrium. There is no learnable equilibrium if both policies are either passive or active. If both policies are active then learnability fails because the equilibrium is explosive.

Comparing our results to the previous literature we find that under passive fiscal policy the conditions for determinacy and learnability on the parameters of the Taylor rule are identical to those found by Bullard and Mitra (2002). More interestingly fiscalist equilibria may well be learnable if monetary policy is passive and fiscal policy is active. This is in contrast to the claim of McCallum (2001), but in line with the results of Evans and Honkapohja (2004). However, unlike Evans and Honkapohja (2004) we find that there are no learnable equilibria if both policies are active. Therefore the monetary and fiscal authority need to coordinate their policies if they want to implement a learnable equilibrium.

The chapter is organised as follows. Section 1.2 introduces the model and section 1.3 sketches the methodology. Section 1.4 reviews the determinacy and learnability properties of the model assuming that fiscal policy balances the budget every period and no government debt is held in equilibrium. Most results of this section are familiar from Bullard and Mitra (2002). However we obtain stronger results than Bullard and Mitra (2002) for determinacy in case of interest rate rules with lagged and forward data. The results of this section serve as a benchmark for section 1.5. Section 1.5 is the main contribution of the chapter. It extends the results of Bullard and Mitra (2002) to the case where lump sum taxes follow the aforementioned feedback rule and equilibrium bond holdings are positive. Finally section 1.6 concludes.

1.2. The Model

This section introduces a cashless New Keynesian model with sticky prices. [5] [6] Firms are monopolistically competitive and prices are adjusted according to Calvo (1983) staggered price setting. The government consists of a monetary and a fiscal authority. Monetary policy is implemented via one of three feedback rules responding to lagged, contemporaneous or expected future variables. Fiscal policy is conducted according to a feedback rule that reacts to outstanding government liabilities as originally proposed by Leeper (1991).

Note that nominal variables will be denoted by upper case letters and real variables by lower case letters. Denoting the aggregate price index by P_t, a nominal variable V_t and the corresponding real variable v_t are related via $v_t = V_t/P_t$. Logarithmic deviations from steady states will be marked with a hat.

[5] In a cashless economy without monetary frictions the interest rate is set by varying the yield on the monetary base and not by open market operations. Arbitrage assures that other riskless assets need to pay the same yield. Although this means that there must be a positive monetary base even in a cashless economy we do not have to distinguish it from other government liabilities like riskless nominal bonds, so that no separate state variable for money will be needed. A detailed discussion is given by Woodford (2003b) chapter 2.1.

[6] We have computed the results for a calibrated economy with money introduced via the utility function and the baseline specification where interest rates are set as a feedback rule of contemporaneous data. The results are undistinguishable from the case of a cashless economy.

1.2.1. Households. The typical household maximises his expected discounted lifetime utility

$$E_0\Big[\sum_{t=0}^{\infty}\beta^t u(c_t,l_t)\Big], \quad \beta\in(0,1), \tag{1}$$

where E_0 is the conditional expectation as of time 0, β is the rate of time preference and $u(\cdot)$ is the per period utility given by

$$u(c_t,l_t)=\frac{c_t^{1-\sigma_c}-1}{1-\sigma_c}-\gamma\frac{l_t^{1+\sigma_l}+1}{1+\sigma_l}, \quad \gamma\geq 0. \tag{2}$$

The household derives utility from consumption c_t and leisure $1-l_t$, where $l_t\in(0,1)$ denotes hours worked. The utility function is assumed to be additively separable in both arguments and of the constant elasticity type. The parameters σ_c and σ_l are the inverse intertemporal elasticities of substitution of consumption and leisure. The maximisation is subject to the budget constraint

$$B_t+P_tc_t+P_t\tau_t\leq P_tw_tl_t+R_{t-1}B_{t-1}+\Omega_t. \tag{3}$$

The household receives income from labour $P_tw_tl_t$, dividends Ω_t from the firms and gross interest payments $R_{t-1}B_{t-1}$ from nominal bond holdings. It pays taxes for $P_t\tau_t$ and consumption P_tc_t and invests in end of period nominal bond holdings B_t. Denoting inflation by $\pi_t=P_t/P_{t-1}$ the budget constraint in real terms is

$$b_t+c_t+\tau_t=w_tl_t+\pi_t^{-1}R_{t-1}b_{t-1}+\Omega_t/P_t. \tag{4}$$

The FOCs and the transversality condition of the household's problem are

$$w_tc_t^{-\sigma_c} = \gamma l_t^{\sigma_l}, \tag{5}$$

$$c_t^{-\sigma_c} = \beta R_tE_t[c_{t+1}^{-\sigma_c}/\pi_{t+1}], \tag{6}$$

$$0 = \lim_{t\to\infty}\beta^t u_{c_t}b_t. \tag{7}$$

1.2.2. Firms. There is a continuum $i\in(0,1)$ of monopolistically competitive firms. Aggregate output y_t is defined as

$$y_t\equiv\Big(\int_0^1 y_{it}^{(\varepsilon-1)/\varepsilon}di\Big)^{\varepsilon/(\varepsilon-1)},$$

where y_{it} is the differentiated good produced by firm i and ε is the elasticity of substitution between any two goods. Cost minimisation yields the goods demand $y_{it}=(P_{it}/P_t)^{-\varepsilon}y_t$ for good i, where

$$P_t=\Big(\int_0^1 P_{it}^{(1-\varepsilon)}di\Big)^{1/(1-\varepsilon)} \tag{8}$$

is the aggregate price index. Output is a linear function of labour input so that $y_{it}=l_{it}$. The pricing of firms is assumed to follow Calvo (1983). Every period each firm adjusts its price only with probability $1-\theta$. A firm i that does not adjust its price sets its price according to $P_{it} = \pi P_{it-1}$, where π is steady state inflation (see Yun, 1996). As demonstrated by Galí (2002), profit maximisation leads to

$$\hat{\pi}_t = \lambda \hat{mc}_t + \beta E_t \hat{\pi}_{t+1}, \tag{9}$$

where $\lambda = (1-\theta)(1-\beta\theta)/\theta > 0$ and mc_t are average real marginal costs. We assume that employment is subsidised at a constant rate ν. Hence marginal costs are given by

$$mc_t = \nu^{-1} w_t. \tag{10}$$

Let $l_t = \int_0^1 l_{it} di$ then it can be shown that $\hat{y}_t = \hat{l}_t$ (cf. Galí, 2002). Combining this equation with (5), (10) and the aggregate resource constraint $y_t = c_t + g_t$ yields

$$\hat{mc}_t = (\sigma_y + \sigma_l)\hat{y}_t - \sigma_y \frac{g}{y} \hat{g}_t, \tag{11}$$

where $\sigma_y \equiv y/c\sigma_c$ is the inverse intertemporal elasticity of substitution of aggregate expenditure. Under flexible prices, a profit maximising firm with an isoelastic demand curve chooses a markup of $\mu = \frac{\varepsilon}{1-\varepsilon}$. It follows that in a flexible price equilibrium marginal costs will be constant and equal to the inverse of the markup, i.e. $mc_t = \mu^{-1}$. Together with (11) this means that potential output under flexible prices $\hat{y}_t^p$ is given by

$$\hat{y}_t^p = \sigma_y/(\sigma_y + \sigma_l) \frac{g}{y} \hat{g}_t. \tag{12}$$

Assume that the employment subsidy exactly offsets the distortion associated with monopolist competition (i.e. $\nu = \mu$). Then the flexible price allocation coincides with the efficient allocation. Define $z_t = y_t - y_t^p$ to be the output gap. Now we have from (11) and (12) that marginal costs are proportionate to the output gap

$$\hat{mc}_t = (\sigma_y + \sigma_l)\hat{z}_t. \tag{13}$$

1.2.3. Government. The government issues nominal bonds B_t and collects taxes $P_t \tau_t$ from the household. It repays bond holdings $R_t B_{t-1}$ of the previous period including interest and finances government expenditures $P_t g_t$. Thus the flow budget constraint of the government is[5]

$$B_t + P_t \tau_t = P_t g_t + R_{t-1} B_{t-1}. \tag{14}$$

Government bonds B_{t+i} are subject to the solvency constraint $\lim_{i\to\infty}(B_{t+i}\prod_{s=0}^{i-1} R_{t+s}^{-1}) = 0$. Following Leeper (1991), lump sum taxes are set according to the tax rule

$$\tau_t = \gamma_0 + \gamma_1 b_{t-1}. \tag{15}$$

We assume throughout that $\gamma_0 \geq g$ and $0 < \gamma_1 < \beta^{-1}$, where g are steady state government expenditures and β^{-1} the gross real interest rate in the steady state. Substituting the tax rule (15) in (14) the budget constraint in real terms is

$$b_t + \gamma_0 + \gamma_1 b_{t-1} = \pi_t^{-1} R_{t-1} b_{t-1} + g_t. \tag{16}$$

In order to assess the consequences for determinacy and learnability monetary policy will be allowed to choose one out of three types of linear feedback rules for the nominal interest rate R_t. Following Bullard and Mitra (2002) we compare lagged, contemporaneous and future data rules

$$\hat{R}_t = \phi_\pi \hat{\pi}_{t-1} + \phi_z \hat{z}_{t-1}, \quad \hat{R}_t = \phi_\pi \hat{\pi}_t + \phi_z \hat{z}_t \quad \text{and} \quad \hat{R}_t = \phi_\pi E_t \hat{\pi}_{t+1} + \phi_z E_t \hat{z}_{t+1}. \tag{17}$$

Throughout this chapter we will assume that ϕ_π and ϕ_z are nonnegative. Finally government expenditures follow the logarithmic AR(1) process

$$\log g_t = \rho \log g_{t-1} + (1-\rho)\log g + \eta_t, \quad \rho \in (0,1), \tag{18}$$

where η_t is white noise.

1.2.4. The Log-Linearised Model. The model can be log-linearised around its non-stochastic steady state. Once interest rates have been set via one of the three types of rules (17) the deviations of output, inflation and bonds from steady state are

$$\hat{\pi}_t = \kappa \hat{z}_t + \beta E_t \hat{\pi}_{t+1}, \tag{19}$$

$$\hat{z}_t = E_t \hat{z}_{t+1} - \sigma_y^{-1}(\hat{R}_t - E_t \hat{\pi}_{t+1}) + (1-\rho)\left(1 + \frac{\sigma_y}{\sigma_y + \sigma_l}\right)\frac{g}{y}\hat{g}_t, \tag{20}$$

$$\hat{b}_t = (\beta^{-1} - \gamma_1)\hat{b}_{t-1} - \beta^{-1}(\hat{\pi}_t - \hat{R}_{t-1}) + \frac{g}{b}\hat{g}_t, \tag{21}$$

where $\kappa \equiv \lambda(\sigma_y + \sigma_l)$. Government expenditures follow the AR(1) process $\hat{g}_t = \rho \hat{g}_{t-1} + \varepsilon_t$. Equation (19) is a New Keynesian Philips curve and (20) is an intertemporal IS curve. Finally (21) is the log-linear approximation of the government budget constraint (16) for a cashless economy. It will be assumed that κ and σ_y are positive and that $0 < \beta < 1$.

Note that equations (17), (19) and (20) can be interpreted as a model where fiscal policy is extremely passive and balances the budget every period. We will subsequently call this model the 'no bonds' case, in contrast to the full model including equation (21), which will be labelled as the case of 'positive bonds'.

1.2.5. Calibration. In order to assure comparability the parameters in the benchmark without bonds are set as in Bullard and Mitra (2002). The discount factor β is set to 0.99. The inverse intertemporal elasticity of substitution σ_y is set to 0.157 and the compound parameter $\kappa = \lambda(\sigma_y + \sigma_l)$ is set to 0.024. Both parameters are estimates from Rotemberg and Woodford (1998) and Rotemberg and Woodford (1999). Finally we take the estimate of $\rho = 0.95$ for the AR(1) coefficient in government expenditures from Galí (2002).

1.3. Methodology

For the analysis of determinacy the model is written in first order form

$$Ax_t = Bx_{t+1} + Cu_t + D\eta_{t+1}, \tag{22}$$

where x_t is a vector of endogenous variables, u_t is a vector of exogenous variables and $\eta_{t+1} = x_{t+1} - E_t x_{t+1}$ is the vector collecting the forecast errors. The model is determinate if there is a unique non-explosive solution, indeterminate if there are multiple non-explosive solutions and explosive otherwise. [7] According to Blanchard and Kahn (1980) the model is (locally) determinate if the number of eigenvalues of $A^{-1}B$ inside the unit circle equals the number of jump variables, it is (locally) indeterminate if there are fewer eigenvalues inside the unit circle than jump variables and explosive otherwise . A jump variable is a variable that is not pinned down by an initial condition. Therefore additional restrictions are needed to identify a unique solution. If there are explosive roots in the system, then some endogenous variables will explode unless a certain linear restriction (a non-explosiveness condition) holds. Provided a specific regularity condition holds the endogenous variables are uniquely identified if the sum of initial conditions and non-explosiveness conditions is equal to the number of endogenous variables. [8] Following the literature we will call a variable that is restricted by an initial condition a predetermined variable. As noted by Sims (2001) the number of jump variables can be read of the system (22) and is equal to the rank of the matrix D. [9]

Regarding learnability we follow the methodology introduced by Marcet and Sargent (1989b) and further developed by Evans and Honkapohja (2001). Compared to the rational expectations benchmark, agents do not know the distribution of the variables they need to forecast. Instead they specify a perceived law of motion for the economy, which is

[7] One may argue for an alternative definition of determinacy especially in case of government bonds, which, by transversality condition (7), are permitted to grow exponentially at a rate of β. However, in this case the local analysis (especially the linearisation) would no longer be valid and so we adapt the more restrictive definition that requires all quantities to be stationary (cf. Woodford, 2003a, p.314 footnote 76).

[8] By regularity condition we mean the rank condition given in Blanchard and Kahn (1980), which assures that each explosive root can be associated with a jump variable. If this condition does not hold, it may e.g. happen that there is an explosive root in an equation where there are only predetermined variables. Then the system will explode even if the number of jump variables equals the number of explosive roots.

[9] As a technical remark note that all theorems in this chapter consider only generic cases. In case of determinacy this amounts to the assumption that none of the characteristic polynomials is self-reciprocal. One special case of self-reciprocity is that roots lie on the unit circle (cf. appendix A.1.2.1 and Henrici (1988) p.496).

continuously adapted to the data via recursive least squares. Thus each time agents receive a new piece of information, they update their PLM and use it to form their expectations. The expectations in turn feed back into the system and change the law of motion of the economy. The question that arises is under which conditions convergence occurs and where this process converges to.

Formally we follow Evans and Honkapohja (2001) and rewrite the model in second order form as

$$x_t = \Gamma_0 + \Gamma_1 E_t x_{t+1} + \Gamma_2 x_{t-1} + \Gamma_3 u_t, \tag{23}$$

where x_t is a vector of endogenous variables, $u_t = \varphi u_{t-1} + \varepsilon_t$ is a vector of exogenous variables which are assumed to follow a stationary VAR. Thus ε_t is white noise and all roots of φ are inside the unit circle. A big issue in the learning literature is the information set that the expectation in equation (23) is conditioned on. There are three standard alternatives. The perhaps most obvious is to use all information on endogenous and exogenous variables as of time t, i.e. to define $E_t x_t := E[x_t | I_t]$ where $I_t = \{x_s, u_s, \varepsilon_s\}_{s=0}^{t}$ is the information available at time t (called t-dating). A widespread alternative is to assume that agents only have information as of time $t-1$, i.e. replace $E_t x_{t+1}$ in (23) by $E_{t-1} x_{t+1}$ (called $t-1$-dating). This approach is popular in the learning literature because it 'avoids a simultaneity between expectations and current values of endogenous variables which may seem more natural in the context of the analysis of learning (see Evans and Honkapohja, 2001, p.229)'. The drawback of modelling learning this way is that most standard models are microfounded and derived under the assumption of t-dating of information. So the model would have to be derived again under a different assumption on the timing of information. The third alternative is to define $I_t = \{x_s, u_s, \varepsilon_s\}_{s=0}^{t-1} \cup \{u_t, \varepsilon_t\}$. In this case the agents are assumed to possess information up to time t on the exogenous variables but only up to time $t-1$ on the endogenous variables. Obviously this mixed dating assumption avoids the simultaneity and above all it is equivalent to t-dating under rational expectations (cf. Evans and Honkapohja, 2001, p.237). Therefore this is the information assumption we will adopt subsequently.

In order to form expectations agents need to specify a perceived law of motion (PLM) for the variables they wish to forecast. The PLM of the agents is assumed to be of the form

$$x_t = c_0 + c_1 x_{t-1} + c_2 u_t, \tag{24}$$

where c_0, c_1 and c_2 are unrestricted matrices of conformable size. An important fact to remember is that learnability does depend on the choice of the PLM. But if an equilibrium is learnable within a certain class of PLMs then it is also learnable in all nested classes. It may though be the case that an equilibrium that is learnable in a small class of PLMs is not learnable in a larger class that includes this PLM. Finally it is obviously not possible to learn an equilibrium that is not included in the class of PLMs. [10] Accordingly if learnability is used to select among different equilibria, then it is important that they are all included in the PLM. Assuming mixed dating and the PLM (24) forecasts are build according to

$$\begin{aligned} E_t x_{t+1} &= c_0 + c_1 E_t x_t + c_2 E_t u_{t+1} && (25) \\ &= (I + c_1) c_0 + c_1^2 x_{t-1} + (c_1 c_2 + c_2 \varphi) u_t && (26) \end{aligned}$$

and so the actual law of motion is

$$\begin{aligned} x_t &= \big(\Gamma_0 + \Gamma_1 (I + c_1) c_0\big) + \big(\Gamma_2 + \Gamma_1 c_1^2\big) x_{t-1} + \big(\Gamma_1 (c_1 c_2 + c_2 \varphi) + \Gamma_3\big) u_t && (27) \\ &= T(c_0, c_1, c_2)(1, x_{t-1}', u_t')' && (28) \end{aligned}$$

implicitly defining the mapping $T(c_0, c_1, c_2)$ from perceived to actual parameters. Note that the symbol I denotes the identity matrix of conformable size. The fixed points of the mapping $T(\bar{c}_0, \bar{c}_1, \bar{c}_2) = (\bar{c}_0, \bar{c}_1, \bar{c}_2)$ are REE of the form (24) and implicitly defined by [11] [12]

$$\begin{aligned} (I - \Gamma_1 - \Gamma_1 \bar{c}_1) \bar{c}_0 &= \Gamma_0, && (29) \\ (I - \Gamma_1 \bar{c}_1) \bar{c}_1 &= \Gamma_2, && (30) \\ (I - \Gamma_1 \bar{c}_1) \bar{c}_2 - \Gamma_1 \bar{c}_2 \varphi &= \Gamma_3. && (31) \end{aligned}$$

It can be shown that the convergence of recursive least squares is governed by the stability of the following matrix differential equation

$$\frac{d}{d\tau}(c_0, c_1, c_2) = T(c_0, c_1, c_2) - (c_0, c_1, c_2). \quad (32)$$

Define E-stability to be the local asymptotic stability of the differential equation (32). Then recursive least squares converges locally to an REE $(\bar{c}_0, \bar{c}_1, \bar{c}_2)$, provided the solution is E-stable and all roots of $\bar{c}_1$ lie inside the unit circle. The last requirement implies that the equivalence between E-stability and convergence of recursive least squares holds only if the REE is stationary. We will see below that the learning algorithm may diverge if the

[10] As argued by Woodford (2003a) the result of McCallum (2001) that fiscalist solutions are not learnable is due to the fact that McCallum does not allow inflation to depend on fiscal variables and so they are by assumption not learnable.

[11] To show that (29)–(31) are indeed the fixed points of the mapping $T(\cdot)$, note e.g. that $c_0 = \Gamma_0 + \Gamma_1 (I + c_1) c_0$ can be written as $c_0 - \Gamma_1 (I + c_1) c_0 = \Gamma_0$, which is obviously the same as (29). Equations (30) and (31) can be derived similarly.

[12] Equation (30) is a quadratic matrix equation, which in general will have multiple solutions. However, given a specific solution $\bar{c}_1$, equation (31) is a linear matrix equation in $\bar{c}_2$ with the generically unique solution $\mathrm{vec}(c_2) = [I \otimes (I - \Gamma_1 \bar{c}_1) - \varphi' \otimes \Gamma_1]^{-1} \mathrm{vec}(\Gamma_3)$.

REE is nonstationary. The derivatives of the T-map (28) are

$$DT_{c_0}(\bar{c}_0, \bar{c}_1, \bar{c}_2) = \Gamma_1(I + \bar{c}_1), \quad (33)$$

$$DT_{c_1}(\bar{c}_0, \bar{c}_1, \bar{c}_2) = \bar{c}_1' \otimes \Gamma_1 + I \otimes \Gamma_1 \bar{c}_1, \quad (34)$$

$$DT_{c_2}(\bar{c}_0, \bar{c}_1, \bar{c}_2) = \varphi' \otimes \Gamma_1 + I \otimes \Gamma_1 \bar{c}_1. \quad (35)$$

where throughout I denotes an identity matrix of conformable size. Under mixed dating, an REE $(\bar{c}_0, \bar{c}_1, \bar{c}_2)$ is E-stable if the eigenvalues of all derivatives (33)–(35) have real parts smaller than one. The converse is also true provided none of the eigenvalues equals one (for details cf. Evans and Honkapohja, 2001, chapter 10).[13]

Finally note that if $\Gamma_2 \neq 0$ there will be multiple REE because equation (30) is a quadratic matrix equation. In these cases it may be difficult to compute $\bar{c}_1$ and therefore to check the E-stability conditions (33)–(35). In cases where it is analytically untractable we will restrict the analysis of learnability to determinate equilibria, which can be computed using one of the algorithms provided by Sims (2001) or Uhlig (1999). Once the REE is known, it is easy to evaluate the E-stability conditions (33)–(35).

The following two sections will investigate the determinacy and E-stability conditions for the model without bonds and the one with positive bond holdings respectively.

1.4. No Bond Holdings

This section considers the case without bond holdings, where the government budget balances every period and thus fiscal policy is substituted out of the model. Most results are familiar from Bullard and Mitra (2002) and serve as a benchmark for the model with positive bond holdings. In case of forward data rules in the policy rule the determinacy result can be strengthened and in case of lagged data necessary and sufficient conditions are presented for equilibrium uniqueness, where Bullard and Mitra (2002) report only sufficient conditions. The case of a government that balances its budget every period is an extreme instance of passive fiscal policy. As will be seen later the conditions obtained under this policy continue to describe determinacy and E-stability of the monetarist solutions under any passive fiscal policy. As determinacy and E-stability will hinge on whether monetary policy is passive or active, we will now state a formal definition.

DEFINITION 1.4.1. Monetary policy is active if a permanent one percentage point increase in inflation $\hat{\pi}_t$ leads to a permanent increase in the nominal interest rate $\hat{R}_t$ larger than one percentage point. It is passive otherwise.

From the New Keynesian Philips curve (19) it can be seen that a permanent increase in inflation by $\Delta\hat{\pi}$ would go along with a permanent increase in the output gap of $\Delta\hat{z} =$

[13] As the E-stability conditions (33)–(35) do not depend on Γ_3, it is not reported in the following.

$(1-\beta)\kappa^{-1}\Delta\hat{\pi}$. Together both effects would lead to a permanent increase in the nominal interest rate of $(\phi_\pi + (1-\beta)\kappa^{-1}\phi_z)\Delta\hat{\pi}$ via one of the interest rate rules (17). Thus monetary policy is active if and only if

$$\phi_\pi + \frac{1-\beta}{\kappa}\phi_z > 1. \tag{36}$$

This condition is the extension of the well known Taylor principle ($\phi_\pi > 1$) to the New Keynesian model. Note that the traditional condition $\phi_\pi > 1$ is sufficient for (36).

1.4.1. Contemporaneous Data Rules. The model with contemporaneous data interest rate rules can be reduced to a first order system in inflation and the output gap by substituting $\hat{R}_t = \phi_\pi \hat{\pi}_t + \phi_z \hat{z}_t$ in the IS curve (20). Then the model can be written in first order form (22) with $x_t = (\hat{\pi}_t, \hat{z}_t)'$ and

$$A = \begin{bmatrix} 1 & -\kappa \\ \phi_\pi & \sigma_y + \phi_z \end{bmatrix}, \quad B = \begin{bmatrix} \beta & 0 \\ 1 & \sigma_y \end{bmatrix}. \tag{37}$$

Determinacy depends on the eigenvalues of $A^{-1}B$, which is given by

$$A^{-1}B = \frac{1}{\phi_z + \sigma_y + \phi_\pi \kappa} \begin{bmatrix} \kappa + \beta(\phi_z + \sigma_y) & \kappa\sigma_y \\ 1 - \phi_\pi \beta & \sigma_y \end{bmatrix}. \tag{38}$$

Both the output gap and inflation are jump variables and therefore, the system is determinate if and only if both roots of $A^{-1}B$ are inside the unit circle.

PROPOSITION 1.4.2. *Under a contemporaneous data interest rate rule there is a determinate rational expectations equilibrium if and only if*

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1. \tag{39}$$

PROOF. See appendix A.1.3. □

Thus the model with contemporaneous data in the interest rate rule is determinate if and only if monetary policy is active. In order to analyse E-stability the model is casted in second order form, which is given by (23) with $x_t = (\hat{\pi}_t, \hat{z}_t)'$, $u_t = \hat{g}_t$, $\alpha = 0$, $\Gamma_1 = A^{-1}B$, $\Gamma_2 = 0$, $\Gamma_3 = -A^{-1}C$ and $\varphi = \rho$. In the current environment ($\Gamma_2 = 0$) it is natural to focus on solutions that like the MSV solution have $\bar{c}_1 = 0$. [14] E-stability requires the eigenvalues of both $DT_{c_0}(\bar{c}_0, \bar{c}_1, \bar{c}_2)$ and $DT_{c_2}(\bar{c}_0, \bar{c}_1, \bar{c}_2)$ to have real parts smaller one. From (33) and (35) we have $DT_{c_0}(\bar{c}_0, \bar{c}_1, \bar{c}_2) = \Gamma_1 = A^{-1}B$ and $DT_{c_2}(\bar{c}_0, \bar{c}_1, \bar{c}_2) = \varphi' \otimes \Gamma_1 = \rho A^{-1}B$. As $\rho < 1$ this is equivalent to the requirement that the eigenvalues of $A^{-1}B$ have real parts smaller one. Determinacy requires the roots of $A^{-1}B$ to be inside the unit circle, so E-stability is implied by determinacy (in case $\bar{c}_1 = 0$). However, in general (even if $\bar{c}_1 = 0$) the reverse conclusion does not hold.

[14] From equation (30) it is obvious that in the current case with Γ_2 the MSV solution has $\bar{c}_1 = 0$.

PROPOSITION 1.4.3. *Under a contemporaneous data interest rate rule the MSV solution is E-stable if and only if*

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1. \tag{40}$$

PROOF. See appendix A.2. □

Proposition 1.4.2 and 1.4.3 together show that determinacy and learnability are both equivalent to active monetary policy. Thus under rules with contemporaneous data learnability also implies determinacy.

1.4.2. Forward Data Rules. Under forward data interest rate rules we can again substitute the nominal interest rate in the IS curve (20) with $\hat{R}_t = \phi_\pi E_t \hat{\pi}_{t+1} + \phi_z E_t \hat{z}_{t+1}$. The endogenous variables are inflation and the output gap $x_t = (\hat{\pi}_t, \hat{z}_t)'$ and the system matrices of the first order form (22) are

$$A = \begin{bmatrix} 1 & -\kappa \\ 0 & \sigma_y \end{bmatrix}, \quad B = \begin{bmatrix} \beta & 0 \\ 1-\phi_\pi & \sigma_y - \phi_z \end{bmatrix}. \tag{41}$$

For determinacy we are interested in $A^{-1}B$ given by

$$A^{-1}B = \frac{1}{\sigma_y} \begin{bmatrix} \kappa(1-\phi_\pi) + \sigma_y\beta & \kappa(\sigma_y - \phi_z) \\ 1-\phi_\pi & \sigma_y - \phi_z \end{bmatrix}. \tag{42}$$

Output and inflation are jump variables and so both roots of $A^{-1}B$ need to be inside the unit circle for determinacy. [15]

PROPOSITION 1.4.4. *Under a forward data interest rate rule there is a determinate rational expectations equilibrium if and only if*

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1, \tag{43}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z < 2\sigma_y. \tag{44}$$

PROOF. See appendix A.1.3. □

Note that determinacy requires again that monetary policy is active (43), but it also requires that the response to the output gap is not too strong (44). As for plausible parameter values $\kappa(1+\beta)^{-1}$ is very small equation (44) is primarily a restriction on the feedback parameter ϕ_z of the output gap in the interest rate rule. The feedback of interest rates on the output gap will also be important under lagged data rules, but as demonstrated above does not matter in case of contemporaneous data rules. As under interest rate rules with contemporaneous data the second order form is given by (23) with $x_t = (\hat{\pi}_t, \hat{z}_t)'$, $u_t = \hat{g}_t$,

[15] Note that in addition to the conditions stated in the following proposition Bullard and Mitra (2002) require that $\phi_z < \sigma_y(1+\beta^{-1})$. However, as demonstrated in the appendix, this condition is redundant as it is implied by (44) and (43).

$\alpha = 0$, $\Gamma_1 = A^{-1}B$, $\Gamma_2 = 0$, $\Gamma_3 = -A^{-1}C$ and $\varphi = \rho$. Additionally it is again the case that $\bar{c}_1 = 0$, so that determinacy implies learnability. However this time the reverse is not true as can be seen from the following proposition.

PROPOSITION 1.4.5. *Under a forward data interest rate rule the MSV solution is E-stable if and only if*

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1. \tag{45}$$

PROOF. See appendix A.2. □

Proposition 1.4.4 and 1.4.5 show that the MSV solution may be E-stable even under indeterminacy. As under contemporaneous data rules E-stability is equivalent to the condition that monetary policy is active (45). However, to obtain determinacy, forward data rules require in addition that the feedback of nominal interest rates to the output gap may not be too large (44). As noted above it is important to look at determinacy and E-stability jointly because E-stability of one equilibrium does not guarantee that there is no other E-stable solution. In fact forward data rules are a very prominent example of policy rules where there are E-stable sunspots. As shown by Honkapohja and Mitra (2004) and Evans and McGough (2005) there are E-stable sunspot solutions if monetary policy is active (45) and the feedback from the output gap to the nominal interest rate is sufficiently large (i.e. the converse of (44) holds).

1.4.3. Lagged Data Rules. The reduced form (22) under lagged data interest rate rules involves the nominal interest rate, inflation and the output gap $x_t = (\hat{R}_t, \hat{\pi}_t, \hat{z}_t)'$. The system matrices are

$$A = \begin{bmatrix} 0 & \phi_\pi & \phi_z \\ 0 & 1 & -\kappa \\ 1 & 0 & \sigma_y \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 1 & \sigma_y \end{bmatrix}. \tag{46}$$

The eigenvalues that determine determinacy are determined by $A^{-1}B$, where

$$A^{-1}B = \frac{1}{\phi_z + \phi_\pi \kappa} \begin{bmatrix} -\sigma_y & \phi_\pi(\kappa + \sigma_y \beta) + \phi_z & \sigma_y(\phi_z + \phi_\pi \kappa) \\ \kappa & \phi_z \beta & 0 \\ 1 & -\phi_\pi \beta & 0 \end{bmatrix}. \tag{47}$$

Under this regime there are again two jump variables (the output gap and inflation) and one predetermined variable (the nominal interest rate), so that two eigenvalues must be inside and one outside the unit circle for determinacy.

PROPOSITION 1.4.6. *Under a lagged data interest rate rule there is a determinate rational expectations equilibrium if and only if*

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z < 1, \tag{48}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z > 2\sigma_y \tag{49}$$

or

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1, \tag{50}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z < 2\sigma_y. \tag{51}$$

PROOF. See appendix A.1.3. □

To analyse E-stability the model is written in second order form (23). Substituting out nominal interest rates $\hat{R}_t = \phi_\pi \hat{\pi}_t + \phi_z \hat{z}_t$ the endogenous and the exogenous variables are $x_t = (\hat{\pi}_t, \hat{z}_t)$ and $u_t = \hat{g}_t$ respectively. The coefficient matrices are $\alpha = 0$, $\varphi = \rho$ and

$$\Gamma_1 = \begin{bmatrix} \beta + \kappa\sigma_y^{-1} & \kappa \\ \sigma_y^{-1} & 1 \end{bmatrix}, \quad \Gamma_2 = -\sigma_y^{-1}\begin{bmatrix} \kappa\phi_\pi & \kappa\phi_z \\ \phi_\pi & \phi_z \end{bmatrix}. \tag{52}$$

The lags in the interest rate rule lead to $\bar{c}_1 \neq 0$. Therefore determinacy does not imply E-stability. Moreover the conditions for E-stability (33)–(35) depend on the parameters of the rational expectations equilibrium. However, it is intractable to compute the rational expectations equilibrium analytically and so there are no analytical results for E-stability. The numerical results for the calibrated model show that the equilibria in the first determinate region, where monetary policy is passive (48) but the response too the output gap is not too small (49) are E-unstable. In contrast the equilibria in the second region of determinacy, where monetary policy is active (50) and the feedback from the output gap to the nominal interest rate is not too large (51) are E-stable. As mentioned above the results are the same as those for the monetarist solution under passive fiscal policy in the model with positive bond holdings to which we will turn now.

1.5. Positive Bond Holdings

The following section extends the results of Bullard and Mitra (2002) incorporating a fiscal policy rule (15) in the sense of Leeper (1991). It also extends the work of Evans and Honkapohja (2004) by including a dynamic IS curve (20) and investigating the effect of different interest rate rules. In this section determinacy and learnability depend not only on the specification of the monetary policy rule but also on the properties of the fiscal policy rule. It will be especially useful to distinguish between active and passive fiscal policy. The subsequent definition follows Woodford (2003b). [16]

DEFINITION 1.5.1. The fiscal policy rule (15) is passive if it implies that $\{\hat{b}_t\}$, as determined by the flow budget constraint (21), is bounded for all small enough values of the endogenous variables $\{\hat{\pi}_t, \hat{R}_t\}$ and the disturbances $\{\hat{g}_t\}$. A policy that is not passive is called active.

[16] The terms passive and active fiscal policy are introduced by Leeper (1991) and correspond to Riccardian and non-Riccardian fiscal policies in the sense of Woodford (2003b, p.312).

From (21) it is obvious that under the tax rule (15) fiscal policy is passive if

$$\gamma_1 > \beta^{-1} - 1 \tag{53}$$

and active if the reverse inequality holds. Recall that γ_1 is the feedback coefficient of lump sum taxes on government debt in the previous period. If this coefficient is larger than the real interest rate, then a shock to real government liabilities is followed by a reduction of these liabilities in the following period, which is financed via an increase in lump sum taxes.

In the analysis of the previous section, where we have abstracted from fiscal policy, we have implicitly assumed that fiscal policy does not affect the conditions for determinacy and E-stability. By definition 1.5.1 a fiscal policy is passive if it leads to a stationary solution for government debt for all paths of output and inflation. Thus passive fiscal policies do not affect the conditions for determinacy. However it is not clear that the conditions for E-stability also remain unaffected. It will however be shown that a straightforward extension of the solutions in the no bonds case bonds continues to be determinate and E-stable under the same conditions as in the previous section provided fiscal policy is passive. This solution has the property that inflation and output are independent of bonds. Following Evans and Honkapohja (2004) we will classify such an REE as a monetarist solution as opposed to a fiscalist solution.

DEFINITION 1.5.2. A fiscalist solution is a rational expectations equilibrium where inflation $\hat{\pi}_t$ is a function of real bond holdings $\hat{b}_t$. A monetarist solution is a rational expectations equilibrium in which inflation $\hat{\pi}_t$ is independent of real bond holdings $\hat{b}_t$.

The definition captures the fact that active fiscal policy leads to an unstable root in the bond equation (21). Determinacy however requires the solution not to explode. As a consequence of the unstable root under active fiscal policy this is only possible if a linear restriction between inflation, the output gap and bonds holds, i.e. $\hat{\pi}_t + k_1\hat{z}_t + k_2\hat{b}_t + k_3\hat{g}_t = 0$. Thus inflation depends on real government debt, which is the fundamental difference between the fiscalist solutions and the traditional monetarist solutions as analysed by Bullard and Mitra (2002).

1.5.1. Contemporaneous Data Rules. Under contemporaneous data rules with bonds the interest rate rule $\hat{R}_t = \phi_\pi \hat{\pi}_t + \phi_z \hat{z}_t$ can be used to replace the nominal interest rate in the budget constraint (21). The model then reduces to a system (22) involving $x_t = (\hat{\pi}_t, \hat{z}_t, \hat{b}_t)'$. The system matrices are

$$A = \begin{bmatrix} 1 & -\kappa & 0 \\ \phi_\pi & \sigma_y + \phi_z & 0 \\ \beta^{-1}\phi_\pi & \beta^{-1}\phi_z & \beta^{-1} - \gamma_1 \end{bmatrix}, \quad B = \begin{bmatrix} \beta & 0 & 0 \\ 1 & \sigma_y & 0 \\ \beta^{-1} & 0 & 1 \end{bmatrix}. \tag{54}$$

For determinacy we are interested in $A^{-1}B$, which is given by

$$A^{-1}B = \begin{bmatrix} \frac{\kappa+\beta(\phi_z+\sigma_y)}{\phi_z+\sigma_y+\phi_\pi\kappa} & \frac{\kappa\sigma_y}{\phi_z+\sigma_y+\phi_\pi\kappa} & 0 \\ \frac{1-\beta\phi_\pi}{\phi_z+\sigma_y+\phi_\pi\kappa} & \frac{\sigma_y}{\phi_z+\sigma_y+\phi_\pi\kappa} & 0 \\ \psi_1 & \psi_2 & (\beta^{-1}-\gamma_1)^{-1} \end{bmatrix}. \tag{55}$$

where the specific form of the parameters ψ_1 and ψ_2 is omitted as they are irrelevant for determinacy. There are two jump variables and one predetermined variable in x_t. For the model to be determinate two roots of $A^{-1}B$ must be inside and one outside the unit circle.

PROPOSITION 1.5.3. *Under a contemporaneous data interest rate rule there is a determinate rational expectations equilibrium if and only if*

$$\gamma_1 > \beta^{-1} - 1, \tag{56}$$

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1 \tag{57}$$

or

$$\gamma_1 < \beta^{-1} - 1, \tag{58}$$

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z < 1. \tag{59}$$

PROOF. See appendix $A.1.3$. □

Proposition 1.5.3 is the result that is closest to the results of Leeper (1991). It states that the rational expectations equilibrium is determinate under a contemporaneous data rule in one of two cases. Either fiscal policy is passive and monetary policy is active or fiscal policy is active and monetary policy is passive. The above result is the analogue of Leeper's result in a flexible price model for a standard New Keynesian model. It was first obtained by Woodford (1996). Figure 1 shows the determinate, indeterminate and explosive region in the space of the central bank's reaction coefficients (ϕ_π, ϕ_z). In order to analyse E-stability it is convenient to rewrite the model in second order form, which is given by (23) with $x_t = (\hat{\pi}_t, \hat{z}_t, \hat{b}_t)$, $\alpha = 0$, $\varphi = \rho$, $u_t = \hat{g}_t$ and [17]

$$\Gamma_1 = \frac{1}{\phi_z + \sigma + \phi_\pi\kappa} \begin{bmatrix} \kappa + \beta(\phi_z+\sigma_y) & \kappa\sigma_y & 0 \\ 1-\phi_\pi\beta & \sigma_y & 0 \\ -\beta^{-1}\kappa - \phi_z - \sigma_y & -\beta^{-1}\kappa\sigma_y & 0 \end{bmatrix}, \quad \Gamma_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta^{-1}\phi_\pi & \beta^{-1}\phi_z & \beta^{-1}-\gamma_1 \end{bmatrix}. \tag{60}$$

Recall that the E-stability conditions (33)–(35) depend on $\bar{c}_1$. Thus we have to compute the E-stability conditions for every REE $(\bar{c}_0, \bar{c}_1, \bar{c}_2)$ separately. There are seven solutions to equation (30) and thus seven different REEs. In the analysis of learnability we will focus on two of these REE, which yield a determinate equilibrium for some values of the policy parameters $(\gamma_1, \phi_\pi, \phi_y)$. It can be checked numerically that given our calibration they are

[17] Cf. appendix A.1.1 for the details of the derivation.

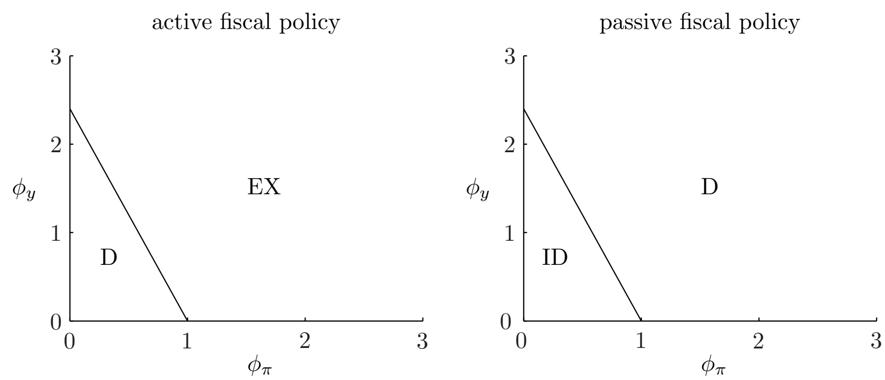

FIGURE 1. *determinate (D), indeterminate (ID) and explosive (EX) region under contemporaneous data rules*

also candidates for an E-stable equilibrium, i.e. are E-stable for some choices of the policy parameters $(\gamma_1, \phi_\pi, \phi_y)$. The remaining five REEs are always E-unstable. The following proposition computes a closed form solution for these two REE. Not surprisingly one is a monetarist and the other one is a fiscalist equilibrium.

PROPOSITION 1.5.4. *Under the conditions* (56) *and* (57) *the determinate solution is the monetarist solution (M) given by* $\bar{c}_0 = 0$, $\bar{c}_1 = \Gamma_2$ *and* $\bar{c}_2$, *where* $\bar{c}_2$ *is the unique solution to* (31).

Under the conditions (58) *and* (59) *the determinate solution is the fiscalist solution (F) given by* $\bar{c}_0 = 0$, $\bar{c}_1 = (\phi_\pi v, \phi_y v, \beta\lambda_3 v)$ *and* $\bar{c}_2$, *where* $\bar{c}_2$ *is the unique solution to* (31) *and* v *is a column vector reported in appendix A.1.4 equation* $(A.41)$.

PROOF. See appendix A.1.4. □

By proposition 1.5.3 there is a determinate solution under passive fiscal policy (56) and active monetary policy (57). Proposition 1.5.4 establishes that it is a specific monetarist solution. There is also a determinate region under active fiscal (58) and passive monetary policy (59) (cf. again proposition 1.5.3). From proposition 1.5.4 we know that the determinate solution under this regime is a certain fiscalist solution. For the monetarist solution it is possible to compute the E-stability conditions analytically.

PROPOSITION 1.5.5. *Under a contemporaneous data interest rate rule the monetarist solution is E-stable if and only if*

$$\gamma_1 > \beta^{-1} - 1, \tag{61}$$

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1 \tag{62}$$

or

$$\gamma_1 < \beta^{-1} - 1, \tag{63}$$

$$\kappa(\phi_\pi - \beta^{-1} + \gamma_1) + \gamma_1 \beta \phi_z > \sigma_y \gamma_1 \beta (\beta^{-1} - \gamma_1 - 1), \tag{64}$$

$$\kappa(2 - \lambda_3)\phi_\pi + (\lambda_3(1-\beta) + 2)\phi_z > \sigma_y(\lambda_3(1+\beta) - 2) + \kappa\beta\lambda_3, \tag{65}$$

where $\lambda_3 = \beta^{-1} - \gamma_1$.

PROOF. See appendix A.2. □

Theorems 1.5.3–1.5.5 show that under passive fiscal policy (61) combined with active monetary policy (62) the monetarist solution is determinate and E-stable. This demonstrates that the results of the previous section (proposition 1.4.3), where fiscal policy is extremely passive, generalise to any passive fiscal policy (see figure 2 on the left). Moreover proposition 1.5.5 states that under active fiscal policy (63) the monetarist solution is E-stable if conditions (64) and (65) hold. In case of the calibrated model shown in figure 2 only condition (64) is binding (cf. the right pane of figure 2). There are two interesting special cases of condition (64). In the borderline case between active and passive fiscal policy, where $\gamma_1 = \beta^{-1} - 1$, it simplifies to the condition that monetary policy is active (62). If on the other hand fiscal policy is extremely active $\gamma_1 = 0$ we have the requirement $\phi_\pi > \beta^{-1}$. Figure 2 (on the right) shows an intermediate case where $0 < \gamma_1 < \beta^{-1} - 1$. Under active fiscal (58) and passive monetary policy (59) the determinate REE is the fis-

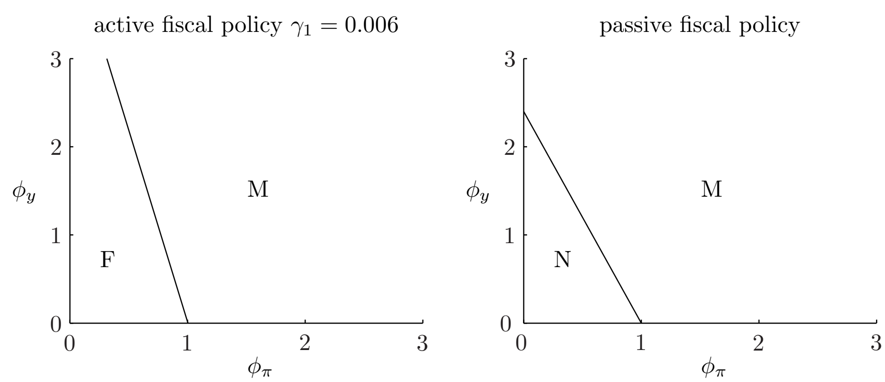

FIGURE 2. *E-stability of the monetarist (M) and fiscalist solution (F), E-instability of all solutions (U) under contemporaneous data rules*

calist solution. The E-stability conditions for the fiscalist solution are too complicated to be solved analytically. The numerical analysis however suggests that the fiscalist solution is E-unstable under passive fiscal policy. Under active fiscal policy the fiscalist solution seems to be E-stable if the converse of condition (64) holds. As mentioned above there is no other E-stable solution in the calibrated model. It is also interesting to note that there is no parameterisation of monetary and fiscal policy where both solutions are E-stable.

Up to now we have only investigated E-stability. It remains to check that the conditions for the equivalence of E-stability and learnability in the sense of convergence of recursive least squares hold. Apart from certain regularity conditions, which do hold in this setup the results in the literature require that the equilibrium that is being learned is stationary. In general E-stability does not guarantee convergence of least squares learning if the equilibrium is explosive. As shown in figure 1 this is the case if both monetary and fiscal policy are active. To check whether convergence still holds we simulate the model for different values of active monetary and fiscal policies. Figure 3 plots the evolution of inflation, the output gap and real government bonds under rational expectations (left panel) and under learning (right panel) for $\gamma_1 = 0.006$, $\phi_\pi = 1.5$ and $\phi_y = 0.5$. The initial values of the beliefs under learning have been chosen close to the monetarist equilibrium. Under rational expectations inflation and the output gap are stationary and only real government bonds are explosive. [18] The fact that agents are learning leads to diverging beliefs under learning. As a consequence inflation and the output gap also diverge as can be seen in the right panel of figure 3. The simulations we have conducted all suggest that neither the monetarist nor the fiscalist solution is stable under learning if both policies are active, even if they are E-stable. Summarising the results for contemporaneous data rules the monetarist solution is the unique learnable and determinate equilibrium for all active monetary and passive fiscal policies. The fiscalist solution is the unique E-stable and determinate equilibrium under passive monetary and active fiscal policy. If both policies are passive the equilibrium is indeterminate and no equilibrium is learnable just as in the orthodox model where fiscal policy simply balances the budget. If both policies are active, all equilibria are explosive. In the monetarist equilibrium inflation and the output gap are stationary and only real bond holdings explode. Our simulations indicate that in this case the all equilibria are unstable under learning even if they are E-stable due to the presence of a non-stationary regressor. [19]

Interest rate rules which depend on current values of the endogenous variables have been criticised not be operational because policy makers usually lack information of real time data. The following two subsections investigate whether these results are robust across two prominent rules that have been proposed in response to this critique.

1.5.2. Forward Data Rules. Under a forward data rule the model with bonds can be reduced to a system involving $x_t = (\hat{\pi}_t, \hat{z}_t, \hat{b}_t)'$ (cf. appendix A.1.1). The system matrices

[18] Yet the transversality condition holds because $\beta(\beta^{-1} - \gamma_1) < 1$ as $0 < \beta < 1$ and $0 < \gamma_1 < \beta^{-1}$.

[19] As noted in section 1.3 E-stability implies learnability under certain regularity assumptions one of which is that the solution is stationary.

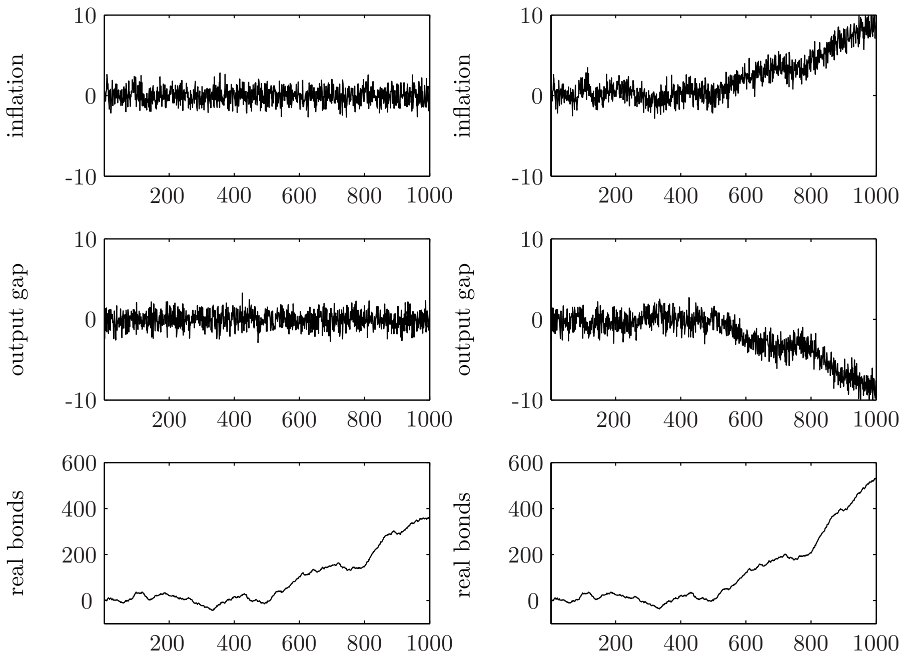

FIGURE 3. *inflation, output gap and real bond holdings when both policies are active and agents have rational expectations (left panel) and are learning (right panel)*

are

$$A = \begin{bmatrix} 1 & -\kappa & 0 \\ 0 & \sigma_y & 0 \\ \beta^{-1}\varphi_\pi & \beta^{-1}\varphi_z & \beta^{-1}-\gamma_1 \end{bmatrix}, \quad B = \begin{bmatrix} \beta & 0 & 0 \\ 1-\phi_\pi & \sigma_y-\phi_z & 0 \\ \beta^{-1} & 0 & 1 \end{bmatrix}. \tag{66}$$

Explicit formulas for φ_π and φ_z are reported in the appendix equation $(A.9)$. For determinacy we are interested in $A^{-1}B$, which is given by [20]

$$A^{-1}B = \begin{bmatrix} \kappa\sigma_y^{-1}(1-\phi_\pi)+\beta & \kappa(1-\sigma_y^{-1}\phi_z) & 0 \\ \sigma_y^{-1}(1-\phi_\pi) & 1-\sigma_y^{-1}\phi_z & 0 \\ \psi_1 & \psi_2 & (\beta^{-1}-\gamma_1)^{-1} \end{bmatrix}. \tag{67}$$

Again for determinacy two roots of $A^{-1}B$ must be inside and one must be outside the unit circle. The proposition below extends the determinacy results in the literature to the case of interest rate rules with forward looking data.

[20]As in the case of a contemporaneous data rule ψ_1 and ψ_2 are compound parameters. Again the determinacy conditions do not depend on them and so they are not reported.

PROPOSITION 1.5.6. *Under a forward data interest rate rule there is a determinate rational expectations equilibrium if and only if either (Case I)*

$$\gamma_1 > \beta^{-1} - 1, \tag{68}$$

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1, \tag{69}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z < 2\sigma_y \tag{70}$$

or (Case II)

$$\gamma_1 < \beta^{-1} - 1 \tag{71}$$

and in addition (Case IIa)

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z < 1, \tag{72}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z < 2\sigma_y \tag{73}$$

or (Case IIb):

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1, \tag{74}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z > 2\sigma_y. \tag{75}$$

PROOF. See appendix A.1.3. □

The first determinacy region (Case I) is the classical one where fiscal policy is passive (68), monetary policy is active (69) and does not respond too aggressively to the output gap (70) (cf. figure 4). The second one (Case IIa) is again similar to the results of Leeper (1991). Active fiscal policy (71) combined with passive monetary policy (71) and – as in the first region – a moderate reaction to the output gap (73) lead to determinacy. However, under interest rate rules with forward data there is a third region of determinacy, where both fiscal and monetary policy are active (71), (74) and monetary policy responds strongly to the output gap (75).

For the analysis of E-stability the model is again transformed into a second order form (23) with $x_t = (\hat{\pi}_t, \hat{z}_t, \hat{b}_t)$, $\alpha = 0$, $\varphi = \begin{bmatrix} \rho & 0 \\ 1 & 0 \end{bmatrix}$, $u_t = (\hat{g}_t, \hat{g}_{t-1})'$ and

$$\Gamma_1 = \begin{bmatrix} \kappa\sigma_y^{-1}(1-\phi_\pi) + \beta & \kappa(1-\sigma_y^{-1}\phi_z) & 0 \\ \sigma_y^{-1}(1-\phi_\pi) & 1-\sigma_y^{-1}\phi_z & 0 \\ -\kappa\sigma_y^{-1}\beta^{-1}(1-\phi_\pi) - 1 & -\kappa\beta^{-1}(1-\sigma_y^{-1}\phi_z) & 0 \end{bmatrix}, \Gamma_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta^{-1}\varphi_\pi & \beta^{-1}\varphi_z & \beta^{-1} - \gamma_1 \end{bmatrix}. \tag{76}$$

The conditions for E-stability depend again on the particular REE under consideration. As with contemporaneous data based rules we concentrate on those REEs which are the determinate solutions under the conditions identified in proposition (1.5.6). It can again be shown numerically for the calibrated economy that any other REE of the form (24) is E-unstable for all configurations of monetary and fiscal policy. The proposition below

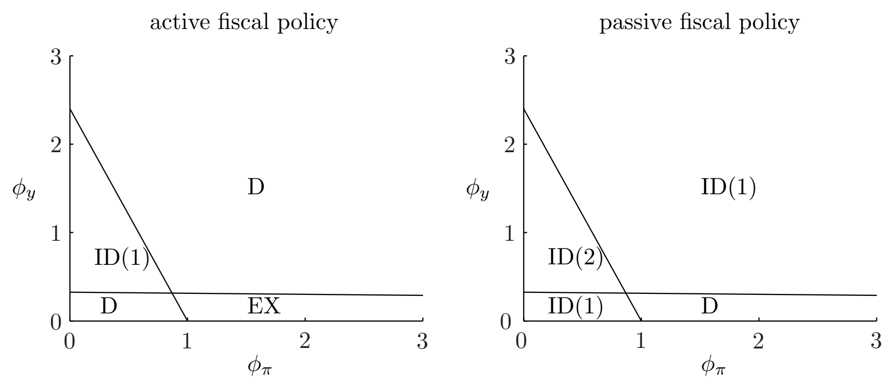

FIGURE 4. *determinate (D), indeterminate (ID) and explosive (EX) region under forward data rules*

establishes that under forward data rules two fiscalist solutions and one monetarist solution are candidates for a determinate equilibrium.

PROPOSITION 1.5.7. *Provided fiscal policy is passive* (68) *and* (69)–(70) *hold the determinate solution is the monetarist solution (M) given by* $\bar{c}_0 = 0$, $\bar{c}_1 = \Gamma_2$ *and* $\bar{c}_2$, *where* $\bar{c}_2$ *is the unique solution to* (31).

Provided fiscal policy is active (71) *there are two fiscalist candidates for a determinate solution, denoted* $(F1)$ *and* $(F2)$. *Under* (72)–(73) *the determinate solution is* $(F1)$ *if* $\phi_z > \sigma_y$ *and* $(F2)$ *if* $\phi_z < \sigma_y$. *Under* (74)–(75) *the determinate solution is* $(F1)$ *if* $\phi_z < \sigma_y$ *and* $(F2)$ *if* $\phi_z > \sigma_y$. *Both fiscalist solutions are of the general form* $\bar{c}_0 = 0$, $\bar{c}_1 = (\varphi_\pi v, \varphi_y v, \beta\lambda_3 v)$ *and* $\bar{c}_2$, *where* $\bar{c}_2$ *is the unique solution to* (31) *and* v *is a column vector reported in appendix A.1.4 equation* $(A.42)$.

PROOF. See appendix A.1.4. □

As in the case of contemporaneous data rules there is a unique monetarist solution (M), which is a candidate for a determinate equilibrium under passive fiscal policy (68). Under active fiscal policy (71) there are two fiscalist candidates (F1) and (F2) for a determinate equilibrium. Which of these solutions is determinate depends on the conditions identified in proposition 1.5.6 and additionally on whether $\phi_z \lesseqgtr \sigma_y$. Again it is possible to derive the conditions for E-stability of the monetarist solution analytically.

PROPOSITION 1.5.8. *Under a forward data interest rate rule the monetarist solution is E-stable if and only if*

$$\gamma_1 > \beta^{-1} - 1, \tag{77}$$

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1 \tag{78}$$

or

$$\gamma_1 < \beta^{-1} - 1, \tag{79}$$

$$\kappa(\phi_\pi - 1) + \gamma_1 \beta \phi_z > \sigma_y \gamma_1 \beta \big(1 - (\beta^{-1} - \gamma_1)^{-1}\big), \tag{80}$$

$$\kappa(\phi_\pi - 1) + \phi_z > \sigma_y \big(\beta + 1 - 2(\beta^{-1} - \gamma_1)^{-1}\big). \tag{81}$$

PROOF. See appendix A.2. □

In case of forward data rules proposition 1.5.8 shows that under passive fiscal policy (77) monetary policy must be active (78) in order to assure E-stability of the monetarist solution (cf. also figure 5) just as in the orthodox model where fiscal policy balances the budget (cmp. proposition 1.4.5). Under active fiscal policy (79) the conditions for E-stability of the monetarist solution are that (80)–(81) hold. As in case of contemporaneous data rules only condition (80) is binding in the calibrated case depicted in figure 5. Another parallel is that in the borderline case between active and passive fiscal policy $\gamma_1 = \beta^{-1} - 1$ the condition simplifies to the requirement that monetary policy be active (78). In case of the extremely active fiscal policy $\gamma_1 = 0$ it simplifies to $\phi_\pi > 1$, which is the traditional version of the Taylor principle.

The conditions for E-stability of the fiscalist solutions are again analytically intractable, so we report only numerical results for the calibrated economy. As in case of contemporaneous data rules the results the fiscalist equilibria are E-unstable if fiscal policy is passive. Under active fiscal policy there is an E-stable fiscalist solution if the converse of condition (80) holds. Depending on the magnitude of the feedback of the output gap on the nominal interest rate ϕ_z, the E-stable solution is given by the fiscalist solution (F1) if $\phi_z > \sigma_y$ and $(F2)$ if $\phi_z < \sigma_y$. As we noted above there is a third region of determinacy under forward data rules, which we have labelled (Case IIb). In this region (the upper right region in figure 4, right panel) the fiscalist equilibrium (F2) is the unique stationary equilibrium, but it turns out to be E-unstable. Instead it is either the other fiscalist solution (F1) or the monetarist solution, which turn out to be E-stable if both policies are active. However, as in case of contemporaneous data rules both types of equilibria are non-stationary. As in case of contemporaneous data rule we have conducted simulations for different parameterisations of active monetary and fiscal policy. They indicate that as a consequence of the non-stationary regressors both are unstable under learning in this parameter region, even though they are E-stable. Summing up the results under forward data rules the conditions for determinacy and learnability under passive fiscal policy are again the same as in the case without bonds (section 1.4). The learnability conditions are also identical to the case of contemporaneous data rules. However, in addition to the case of contemporaneous data rules the magnitude of the feedback coefficient of the output gap on the nominal rate of interest ϕ_z is important for uniqueness of the equilibrium. If fiscal policy is passive and monetary policy is active

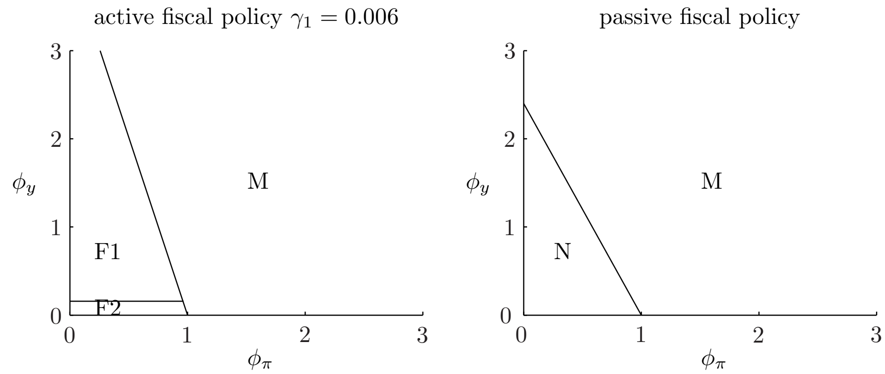

FIGURE 5. *E-stability of the monetarist (M) and fiscalist solutions (F1,F2), E-instability of all solutions (U) under forward data rules*

then the monetarist solution is learnable . If in addition ϕ_z is sufficiently small, then it is the determinate and also the unique learnable equilibrium. For large values of ϕ_z the equilibrium is indeterminate and there may exist learnable sunspot equilibria. If fiscal policy is active and monetary policy is passive then there is a learnable fiscalist solution. Again, if ϕ_z is sufficiently small the fiscalist solution is the determinate and also the unique learnable equilibrium. The equilibrium is indeterminate and unstable under learning if both policies are passive. Finally, all equilibria are explosive and unstable under learning if both policies are active. The main difference compared to contemporaneous data rules is that, independent of what fiscal policy does, monetary policy should choose a coefficient ϕ_z which is not too large to prevent the possibility of learnable sunspots.

1.5.3. Lagged Data Rules. The last type of interest rate rules which we consider are lagged data rules $\hat{R}_t = \phi_\pi \hat{\pi}_{t-1} + \phi_z \hat{z}_{t-1}$. With $x_t = (\hat{R}_t, \hat{\pi}_t, \hat{z}_t, \hat{b}_t)'$ and

$$A = \begin{bmatrix} 0 & \phi_\pi & \phi_z & 0 \\ 0 & 1 & -\kappa & 0 \\ 1 & 0 & \sigma_y & 0 \\ \beta^{-1} & -\beta^{-1} & 0 & \beta^{-1} - \gamma_1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 1 & \sigma_y & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{82}$$

the model can be written in first order form (cf. appendix $A.1.1$). The relevant matrix to investigate determinacy is

$$A^{-1}B = \begin{bmatrix} -\frac{\sigma_y}{\phi_z + \phi_\pi \kappa} & 1 + \frac{\phi_\pi \sigma_y \beta}{\phi_z + \phi_\pi \kappa} & \sigma_y & 0 \\ \frac{\kappa}{\phi_z + \phi_\pi \kappa} & \frac{\phi_z \beta}{\phi_z + \phi_\pi \kappa} & 0 & 0 \\ \frac{1}{\phi_z + \phi_\pi \kappa} & -\frac{\phi_\pi \beta}{\phi_z + \phi_\pi \kappa} & 0 & 0 \\ \psi_1 & \psi_2 & -\frac{\sigma_y}{1 - \gamma_1 \beta} & (\beta^{-1} - \gamma_1)^{-1} \end{bmatrix}. \tag{83}$$

Noting that x_t contains two jump variables and two predetermined variables and that the system is block triangular it is straightforward to derive the conditions for determinacy.

PROPOSITION 1.5.9. *Under a lagged data interest rate rule there is a determinate rational expectations equilibrium if only if (Case I)*

$$\gamma_1 > \beta^{-1} - 1 \tag{84}$$

and either (Case Ia)

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z > 1, \tag{85}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z < 2\sigma_y \tag{86}$$

or (Case Ib)

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z < 1, \tag{87}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z > 2\sigma_y \tag{88}$$

or (Case II)

$$\gamma_1 < \beta^{-1} - 1, \tag{89}$$

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_z < 1, \tag{90}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z < 2\sigma_y. \tag{91}$$

PROOF. See appendix A.1.3. □

In case of lagged data rules there are three regions with a determinate equilibrium. Under passive fiscal policy (84) one set of conditions is again that monetary policy is active (85) and interest rates are not too sensitive to the output gap (86). Under active fiscal policy (89) determinacy requires that monetary policy is passive (90) and does not react too much to the output gap (91). So far the conditions for determinacy are the same as under forward data rules. However under lagged data rules there is a second determinacy region under passive fiscal policy (84) where monetary policy is passive (87) and the sensitivity of interest rates with respect to changes in the output gap is not too small (88). The determinate regions for the calibrated model are depicted in figure 6.

Again we derive a second order form to analyse E-stability. It is given by (23) with $x_t = (\hat{\pi}_t, \hat{z}_t, \hat{R}_t, \hat{b}_t)$, $\alpha = 0$, $\varphi = \rho$, $u_t = \hat{g}_t$ and

$$\Gamma_1 = \begin{bmatrix} \beta + \kappa\sigma_y^{-1} & \kappa & 0 & 0 \\ \sigma_y^{-1} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1-\beta^{-1}\kappa\sigma_y^{-1} & -\beta^{-1}\kappa & 0 & 0 \end{bmatrix}, \quad \Gamma_2 = \begin{bmatrix} -\kappa\phi_\pi\sigma_y^{-1} & -\kappa\phi_z\sigma_y^{-1} & 0 & 0 \\ -\phi_\pi\sigma_y^{-1} & -\phi_z\sigma_y^{-1} & 0 & 0 \\ \phi_\pi & \phi_z & 0 & 0 \\ \beta^{-1}\kappa\phi_\pi\sigma_y^{-1} & \beta^{-1}\kappa\phi_z\sigma_y^{-1} & \beta^{-1} & \beta^{-1}-\gamma_1 \end{bmatrix}. \tag{92}$$

In case of lagged data rules the reduced form of the model is four dimensional and therefore it is not possible to solve for the REEs analytically. Therefore we will restrict the analysis of E-stability to the cases where there is a determinate solution, which can be computed

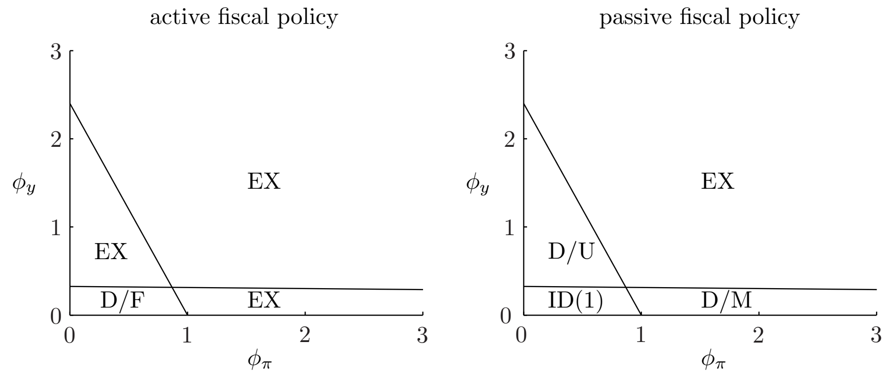

FIGURE 6. *determinacy and E-stability under lagged data rules (determinacy (D), indeterminacy (ID), explosiveness (EX); E-stability of a monetarist (M) or fiscalist (F) solution; E-instability (U))*

numerically. Figure 6 depicts the results for the calibrated model. It turns out that the determinate region where both policies are passive is not E-stable, whereas the regions where one of the policies is active and the other one is passive are both E-stable. It is interesting to note that the regions where there is a determinate and E-stable solution are the same as under forward data rules.

1.6. Conclusions

We have investigated the consequences of policy interaction for determinacy and learnability in a standard New Keynesian model. Following McCallum (2001) and Evans and Honkapohja (2004) we have grouped the different solutions in monetarist and fiscalist solutions. As one may have expected the results differ fundamentally when monetary and fiscal policy are active or passive.

Consider first the case of passive fiscal policy. By definition the conditions for determinacy under passive fiscal policy are the same as in a model where fiscal policy is ignored (as e.g. in Bullard and Mitra, 2002). The solutions identified by Bullard and Mitra (2002) correspond to the monetarist solutions in the model with a fiscal policy rule. Under passive fiscal policy the conditions for E-stability of the monetarist solutions continue to be the same as in Bullard and Mitra (2002). Notably they are independent of the fiscal policy rule. Across all the three types of interest rate rules there is a unique learnable monetarist equilibrium if monetary policy is active and the nominal interest rate is not too sensitive to changes in the output gap. If on the other hand monetary policy is passive then there is no learnable equilibrium. Under passive fiscal policy the fiscalist solutions are all E-unstable.

As mentioned in the introduction there is a dispute in the literature about the plausibility of the predictions of the fiscal theory of the price level under active fiscal policy. The

FTPL has been criticised on various grounds (cf. e.g. Buiter, 2002) and recently the instability under learning has been brought forward against the FTPL by McCallum (2001). Evans and Honkapohja (2004) find that, while fiscalist equilibria are not learnable in the model of McCallum (2001), they may indeed be learnable for certain monetary and fiscal policies in a model similar to the one of Leeper (1991). The New Keynesian model has probably become the most popular model in the analysis of monetary policy. We find significant support for the predictions made by the FTPL in a New Keynesian framework. In the primary case dealt with in the literature on the FTPL, where fiscal policy is active and monetary policy is passive there is a learnable fiscalist equilibrium for a large range of parameterisations of the policy rules. Especially there is a unique learnable fiscalist equilibrium if the nominal interest rate should not vary too much with the output gap, the same condition which guarantees a unique learnable monetarist equilibrium under passive fiscal and active monetary policy. This result is robust across various popular specifications of the interest rate rule. The orthodox monetarist solutions are found to be E-unstable under passive monetary policy.

Finally, all solutions are explosive in a regime where both policies are active. This is important for the analysis of learnability because E-stability guarantees convergence of least squares learning only if the equilibrium that is being learned is stationary. There are no analytical results for the convergence of learning to explosive equilibria, but the simulations we have conducted suggest that neither equilibrium is learnable under this policy regime.

Contrary to a conjecture by McCallum (2001) the theory of adaptive learning shows a reasonable way how – under active fiscal and passive policy – agents may coordinate on a an equilibrium in which fiscal variables influence the evolution of the price level as predicted by the FTPL. Yet there is still a unique learnable monetarist equilibrium if monetary policy is active and fiscal policy is passive. If both policies are either passive or active then there is no learnable equilibrium. These results suggest that the coordination of monetary and fiscal policy is an important ingredient to achieve a unique learnable equilibrium.

CHAPTER 2

Convergence of Least Squares and E-Stability in Models with Sunspots

2.1. Introduction

Adaptive learning has become increasingly popular over the last two decades. The concept of E-stability has been shown to yield the conditions for learnability of fundamental equilibria in various models. Recently a growing literature on the learnability of non-fundamental equilibria has emerged. The relevance of non-fundamental equilibria for economic analysis is a contentious issue because the concept of rational expectations equilibria (REE) does not reveal how the coordination on a particular equilibrium takes places. Least squares learning is a widely applied adaptive learning algorithm, which shows how the coordination can take place and which can be used as an equilibrium selection device. In case of fundamental equilibria it has been shown that E-stability is equivalent to convergence of least squares learning to an REE (cf. Evans and Honkapohja, 2001). For non-fundamental equilibria a connection between the convergence of learning algorithms like recursive least squares (RLS) and E-stability remains to be proven. Yet there is a growing literature on the learnability of sunspot equilibria that focuses entirely on E-stability. Prominent examples are Honkapohja and Mitra (2004) and Evans and McGough (2005) who investigate E-stability of sunspot equilibria in models of monetary policy. Indeed both papers only check certain conditions, which are conjectured to be necessary and sufficient for E-stability, but so far this has not been proven either.

The main contributions of this chapter are that we present new results on E-stability and the convergence of least squares, which have not been established before for sunspot equilibria. For common factor representations we give necessary and sufficient conditions for E-stability. We also show that general form representations are always E-unstable in models purely forward looking models. This result is likely to extend to any model with a sufficiently low degree of inertia. Apart from giving necessary and sufficient conditions for E-stability we also establish that the convergence of least squares learning (augmented with projection facilities) is governed by the stability of the mean ODE. We conjecture that E-stability of a set of sunspot representations is equivalent to stability of the mean ODE analogous to the results for fundamental equilibria. However, we have not yet been able to prove this formally. For common factor representations we present evidence from a simulation study that indeed stability under least squares learning E-stability are equivalent.[1]

The rest of the chapter is structured as follows. Section 2.2 sets up a multivariate self-referential linear model and defines a rational expectations equilibrium, an equilibrium representation and E-stability. The central part of the section are two theorems, one that gives necessary and sufficient conditions for E-stability of common factor representations and another one which proves that general form representations are always E-unstable in purely forward looking models. In a purely forward looking model common factor representations are E-stable provided the MSV representation is E-stable. Section 2.3 establishes that under standard regularity conditions recursive least squares (RLS) converges locally to an REE provided the mean ODE is locally stable. We present a convergence result that uses projection facilities to assure that the learning algorithm is well behaved.[2] The convergence proof is based on a general convergence theorem by Delyon (1996). Section 2.4 applies the results to a standard New Keynesian model where interest rates are set via a forward looking Taylor rule. Evans and McGough (2005) claim that there are E-stable common factor sunspot representations. As mentioned above all results in the literature are based on the conjecture that the conditions for E-stability are analogous to those for fundamental equilibria. We can apply our results on E-stability of common factor representations from section 2.2 which confirm that this conjecture is indeed correct. In section 2.4 we present some simulation results which show convergence of RLS to a common factor sunspot representation even without imposition of a projection facility. On the other hand a failure of E-stability leads to rapid divergence of the beliefs. These results show that the equivalence of E-stability and convergence of least squares learning, which has been established for fundamental equilibria, seems to carry over to the case of non-fundamental equilibria. Another interesting outcome of the simulations is that under a constant gain version of least squares learning there may be large fluctuations in the beliefs of the agents about the impact of the sunspot on the economy. In contrast to the results under RLS the constant gain algorithm does not settle down on a particular equilibrium. It rather wanders around in the continuum of sunspot representations, so under learning there may be times with a large and other times with a small impact of sunspots on the economy.

[1] Our instability result explains why general form representations have been found to be E-unstable for reasonable parameter values in all models we are aware of. The lack of E-stable representations renders the study of a connection between E-stability and converges of RLS impossible in case of general form representations. Therefore, in our simulation study, we restrict our attention to common factor representations.

[2] The projections are generally needed to assure that the estimates remain in a compact set and that the learning algorithm converges. For details confer section 2.3.

2.2. A Multivariate Linear Model

Consider the multivariate linear model

$$
\begin{aligned}
y_n &= \alpha + \beta E_n^* y_{n+1} + \delta y_{n-1} + \kappa w_n, && (93)\\
w_n &= \rho w_{n-1} + \vartheta_n, && (94)
\end{aligned}
$$

where ϑ_n is white noise and the roots of ρ are inside the unit circle. The symbol E^* denotes the expectation operator. The star is meant to indicate that expectations need not be rational. Assume that at time n the information set is given by $I_n = \{y_{k-1}, w_k, \varepsilon_k; k \leq n\}$, where ε_n is an arbitrary martingale difference sequence. So all endogenous and exogenous variables are observable, but the information set in period n excludes the current endogenous variables. Under rational expectations this information set is equivalent to the one which includes current endogenous variables, because the current endogenous variables y_n can be inferred from the model (93)–(94) given I_n. Under learning the equivalence breaks down. The standard approach in the learning literature is to assume that current endogenous variables realise after the agents have formed their expectations about y_{n+1} and thus not to include them in the information set of the agents in period n.

DEFINITION 2.2.1. A *rational expectations equilibrium* (REE) is a stochastic process y_n that solves the expectational difference equation (93)–(94) with $E_n^* = E_n$, where E_n denotes the conditional expectation operator as of time n.

DEFINITION 2.2.2. A *rational expectations equilibrium representation* (REER) is an explicit difference equation any solution to which is a rational expectations equilibrium.

An REER is a recursive representation of an REE without a reference to expectations. In the learning literature it is usually computed using the method of undetermined coefficients. An REE may have multiple recursive representations. The minimal state variable (MSV) representation has become the standard representation for regular REEs (cf. McCallum, 1983).

DEFINITION 2.2.3. A *minimal state variable representation* (MSV representation) of the model (93)–(94) is an REER of the type

$$y_n = a + b y_{n-1} + c w_n, \tag{95}$$

with $b = 0$ if $\delta = 0$.

Definition 2.2.3 is the standard definition used in the learning literature. In a model with lags ($\delta \neq 0$) there may be multiple representations of this form.[3] In the presence of sunspots there are two prominent types of representations for the same REE. The learnability of an REE depends on the particular representation of the equilibrium. Therefore it is important to distinguish between the two concepts in the analysis of learnability.

DEFINITION 2.2.4. Let ε_n is an arbitrary martingale difference sequence. A *general form representation* of the model (93)–(94) is an REER of the type

$$y_n = a + by_{n-1} + hy_{n-2} + cw_n + fw_{n-1} + e\varepsilon_n, \tag{96}$$

with $h = 0$ and $f = 0$ if δ=0.

Note that the coefficients of an REER have to meet certain conditions, which will be laid out below. So not every process of the functional form (96) is a general form representation. General form representations have often been found to be unstable under learning. However, 'every sunspot equilibrium has a common factor representation, and the stability properties of common factor representations are different from their general form counterparts' (Evans and McGough, 2005).

DEFINITION 2.2.5. Let ε_n is an arbitrary martingale difference sequence. A *common factor representation* of the model (93)–(94) is an REER of the type

$$y_n = a + by_{n-1} + cw_n + d\zeta_n, \tag{97}$$

$$\zeta_n = \varphi\zeta_{n-1} + \varepsilon_n \tag{98}$$

with $b = 0$ if $\delta = 0$.

In case of indeterminacy of order k, ζ_n is a k dimensional vector of sunspots and φ is a k dimensional diagonal matrix, with the stable characteristic roots of the model on the diagonal (cf. Evans and McGough, 2005). Therefore ζ_n is often called a resonant frequency sunspot. Compared to general form representations, common factor representations have more often been found to be E-stable. The fact that general form representations include additional lags of the endogenous variables as compared to common factor representations seems to be destabilising under learning. The additional lags are also the driving force of the instability result for general form representations, which we will present below.

[3]Learnability is a criterion to chose among multiple REERs. Therefore the literature on adaptive learning usually does not use the selection criterion proposed by McCallum (1983) to choose a particular REER. Especially all REERs of the form (95) are labelled MSV representation and not only those selected by McCallum's subsidiary principle.

Because of its relative simplicity assume first that the perceived law of motion (PLM) of the agents is of the same functional form as a common factor representation (97).[4] Recall that the current endogenous variables y_n are not included in the information set I_n and assume that ρ is known. Then the forecasts at time n are

$$\begin{aligned} E_n^* y_{n+1} &= a + bE_n^* y_n + cE_n^* w_{n+1} + dE_n^* \zeta_{n+1} \\ &= a + b(a + by_{n-1} + cw_n + d\zeta_n) + c\rho w_n + d\varphi\zeta_n \\ &= (I+b)a + b^2 y_{n-1} + (bc + c\rho)w_n + (bd + d\varphi)\zeta_n. \end{aligned}$$

The star indicates that during the learning process the PLM of the agents will not be equal to an REER and as a consequence their forecasts $E_n^* y_{n+1}$ will deviate from rational forecasts.[5] If the REER is learnable, then the error made by the agents will vanish eventually. Inserting these expectations in (93) and collecting coefficients yields the T-map

$$a \rightarrow T_a(a,b) = \alpha + \beta(I+b)a, \tag{99}$$

$$b \rightarrow T_b(b) = \beta b^2 + \delta, \tag{100}$$

$$c \rightarrow T_c(b,c) = \beta(bc + c\rho) + \kappa, \tag{101}$$

$$d \rightarrow T_d(b,d) = \beta(bd + d\varphi). \tag{102}$$

If the PLM is of the general form (96) and assuming ρ is known the agents' forecasts are

$$\begin{aligned} E_n^* y_{n+1} &= a + bE_n^* y_n + hy_{n-1} + cE_n^* w_{n+1} + fw_n \\ &= a + b(a + by_{n-1} + hy_{n-2} + cw_n + fw_{n-1} + e\varepsilon_n) + hy_{n-1} + c\rho w_n + fw_n \\ &= (I+b)a + (b^2 + h)y_{n-1} + bhy_{n-2} + (bc + c\rho + f)w_n + bfw_{n-1} + be\varepsilon_n. \end{aligned}$$

Again, inserting these expectations in (93) yields the T-map given by equation (99) together with

$$(b,h) \rightarrow T_{(b,h)}(b,h) = (\beta(b^2+h) + \delta, \beta bh), \tag{103}$$

$$c \rightarrow T_c(b,c,f) = \beta(bc + c\rho + f) + \kappa, \tag{104}$$

$$f \rightarrow T_f(b,f) = \beta bf, \tag{105}$$

$$e \rightarrow T_e(b,e) = \beta be. \tag{106}$$

In case the perceived law of motion is of the same functional form as a common factor representations define $\xi' = (a,b,c,d)$ and $z_n' = (1, y_{n-1}', w_n', \zeta_n')$. If it is of the general form

[4] That is the PLM is an arbitrary process of the form (96). Especially all coefficients are unrestricted matrices.

[5] The literature on adaptive learning generally assumes that the agents ignore the uncertainty of the coefficients during the learning process, when they make their forecasts.

let $\xi' = (a, b, h, c, f, e)$ and $z_n' = (1, y_{n-1}', y_{n-2}', w_n', w_{n-1}', \varepsilon_n')$. With these definitions we can write the perceived law of motion in the simple form

$$y_n = \xi' z_n. \tag{107}$$

The counterpart of the perceived law of motion (PLM) is the actual law of motion (ALM). The PLM is the forecasting model of the agents and the ALM is the resulting data generating process for the economy, when the PLM is substituted in the model (93)–(94). The coefficients of the ALM are given by the T-map. In case of a PLM of the common factor type let $T(\xi) = (T_a(a,b), T_b(b), T_c(b,c), T_d(b,d))$, where for all $v \in \{a,b,c,d\}$ the mappings $T_v(\cdot)$ are defined by (99)–(102). In case of a PLM of the general form define the T-map analogously with the components $T_v(\cdot)$, $v \in \{a,(b,h),c,f,e\}$ given by (99) and (103)–(106). Then the actual law of motion (ALM) is

$$y_n = T(\xi)' z_n. \tag{108}$$

The T-map $\xi \to T(\xi)$ is a convenient representation of the results of the method of undetermined coefficients. A fixed point $\bar{\xi} = T(\bar{\xi})$ of this mapping yields the coefficients $\bar{\xi}$ of the corresponding REER $y_n = \bar{\xi}' z_n$. The following proposition collects these results.

PROPOSITION 2.2.6. *A common factor representation is a difference equation of the form* (97)*, where the coefficients* (a,b,c,d) *are a fixed point of the T-map given by* (99)–(102)*. A general form representation is a difference equation of the form* (96)*, where the coefficients* (a,b,h,c,f,e) *are a fixed point of the T-map given by* (99) *and* (103)–(106).

Apart from its importance for the computation of an REER the T-map plays a prominent role in the literature on adaptive learning. Across many different models it has turned out that the T-map determines the conditions for learnability of an REER under various learning algorithms, especially RLS. The central property of the T-map in this respect is *expectational stability* (E-stability). In case of a locally determinate equilibrium E-stability is usually defined as local asymptotic stability of the REER $\bar{\xi}$ under the differential equation

$$\dot{\xi} = T(\xi) - \xi \tag{109}$$

where the T-map maps the coefficients of the perceived law of motion of the agents into those of the actual law of motion. By definition the REER $\bar{\xi}$ is a fixed point of the T-map and thus also a rest point of the ODE (109). Under certain regularity conditions E-stability guarantees that least squares learning converges locally to $\bar{\xi}$. [6] This definition of E-stability is no longer useful if the REER is not locally unique as it is the case for a sunspot representation. In general there will a continuum of common factor or general form representations $\bar{\xi}$. In case of indeterminacy there is a continuum of sunspot representations

[6] The regularity conditions will be introduced in section 2.3.

and we cannot expect convergence to a particular equilibrium. A suitable extension of the concept of learnability is to ask whether convergence occurs to some point in the set of sunspot representations Ξ for all initial conditions in a neighbourhood of the set.

DEFINITION 2.2.7. Let Ξ be a set of common factor or general form representations of the model (93)–(94). Then Ξ is *E-stable* if it is a locally asymptotically stable set in the sense of Lyapunov for equation (109), where the T-map is the same as in proposition 2.2.6.

Definition 2.2.7 is based on the definition by Evans and Honkapohja (2001) p.245.[7] It generalises the definition of E-stability of an MSV representation by not requiring convergence to a particular equilibrium point $\bar{\xi}$ but only to a set of equilibrium points Ξ.[8] We will now apply this concept to both common factor and general form representations.

Consider first the case of common factor representations. Generically there are $\binom{2k}{k}$ isolated solutions for $\bar{b}$, where k is the number of endogenous variables (i.e. the dimension of y_n). Let $\Xi_{\bar{b}} = \{\xi | T(\xi) = \xi \wedge b = \bar{b}\}$ be the set of REERs corresponding to one particular solution to equation (100). In the absence of sunspots $\Xi_{\bar{b}}$ contains a single REER. This is because generically a and c are uniquely determined by (99) and (101) given $\bar{b}$. Moreover absence of sunspots means that $\bar{d} = 0$ yielding a unique REER $\bar{\xi}$ of the MSV type. However, if sunspots exist then there will be a continuum of solutions for $\bar{d}$. In this case $\Xi_{\bar{b}}$ is a non-trivial set of REERs. Before we present the conditions for E-stability define the following derivatives

$$\begin{aligned} DT_a(\bar{b}) &= \beta(I+\bar{b}), & (110)\\ DT_b(\bar{b}) &= \bar{b}' \otimes \beta + I \otimes \beta\bar{b}, & (111)\\ DT_c(\bar{b}) &= \rho' \otimes \beta + I \otimes \beta\bar{b}, & (112)\\ DT_d(\bar{b}) &= \varphi \otimes \beta + I \otimes \beta\bar{b}, & (113) \end{aligned}$$

where for all $v \in \{a,b,c,f\}$, $DT_v = \partial \text{vec } T_v / \partial (\text{vec } v)'$ is the Jacobian of vec T_v.

PROPOSITION 2.2.8. *Assume that the information set at time n is* $I_n = \{1, y_{n-1}, w_n, \varepsilon_n\}$. *Let* $m = \dim(\{x | (DT_d(\bar{b}) - I)x = 0\})$. *A necessary condition for E-stability of a set of common factor representations* $\Xi_{\bar{b}}$ *is that none of the eigenvalues of* $DT_a(\bar{b})$, $DT_b(\bar{b})$, $DT_c(\bar{b})$ *and* $DT_d(\bar{b})$ *has a real part larger than 1. This condition is also sufficient if apart from m eigenvalues of* $DT_d(\bar{b})$ *equal to 1 the eigenvalues of* $DT_a(\bar{b})$, $DT_b(\bar{b})$, $DT_c(\bar{b})$ *and* $DT_d(\bar{b})$ *have real parts different from 1.*

[7] A formal definition of local asymptotic stability in the sense of Lyapunov is given in the appendix (cf. definition A.2.1).

[8] In our simulations in section 2.4 we will see that the particular point of convergence depends both on the initial beliefs ξ_0 and on the shocks that hit the economy.

PROOF. See appendix A.2. □

Proposition 2.2.8 is the main proposition of this section. Lets make the regularity assumption that apart from m eigenvalues of the Jacobian of the T-map equal to one, there are no further eigenvalues equal to one. Then proposition 2.2.8 states that the set of REERs $\Xi_{\bar{b}}$ is E-stable if and only if all remaining eigenvalues have real parts smaller than one. The conditions on the matrices (110)–(113) are well known in the literature. However, due to the zero eigenvalues in $DT_d(\bar{b})$ standard stability results are not applicable and to our knowledge this is the first time that it is proven that they are indeed necessary and sufficient for E-stability.

Inspection of the T-map for common factor representations (equations 99–102) shows that the components $T_a(\cdot)$, $T_b(\cdot)$ and $T_c(\cdot)$ are independent of the sunspot multiplier d. Thus a fixed point $(\bar{a}, \bar{b}, \bar{c})$ yields an MSV representation. Put differently, a common factor representation with $\bar{d} = 0$ is an MSV representation. Therefore a common factor representation is E-stable if the corresponding MSV representation (the one with the same $\bar{b}$) is E-stable and none of the eigenvalues of $DT_d(\bar{b})$ has a real part larger than 1. The condition on the eigenvalues of $DT_d(\bar{b})$ does not seem to be very restrictive in applications, which explains why common factor representations are often found to be E-stable, whenever the MSV solution is E-stable. For the special case of purely forward looking models we have the even stronger result that E-stability of the set of common factor representations and the MSV representation are equivalent.

COROLLARY 2.2.9. *Assume that the model is purely forward looking ($\delta = 0$) and that sunspots exist. Then E-stability of the MSV representation is equivalent to E-stability of the set of common factor sunspot representations. The necessary and sufficient condition is that all eigenvalues of β have real parts smaller than 1.*

In a purely forward looking model $DT_d(\bar{b}) = \varphi \otimes \beta$. The result now follows because the eigenvalues of the Kronecker product are equal to the products of the eigenvalues of the matrices and the eigenvalues of φ are by construction stable eigenvalues of the model.[9] A prominent application of this framework to monetary policy is the New Keynesian model with a forward looking interest rate rule (cf. e.g. Evans and McGough, 2005). We will review this framework in section 2.4. Corollary 2.2.9 is a bit more general than it seems at first sight. The eigenvalues of a matrix are continuous functions of the parameters of the model. It follows immediately that corollary 2.2.9 extends to all models with a sufficiently low degree of inertia.

[9] See Evans and McGough (2005) for details concerning the computation of general form representations.

In case of general form representations define the set $\Xi_{(\bar{b},\bar{h})} = \{\xi | T(\xi) = \xi \wedge (b,h) = (\bar{b},\bar{h})\}$, where $(\bar{b},\bar{h})$ is a particular solution to (103). As in case of common factor representations $\Xi_{(\bar{b},\bar{h})}$ may contain a single REER or a continuum of REERs. The derivatives of the T-map are (110) and (112) as well as

$$DT_{(b,h)}(\bar{b},\bar{h}) = \begin{bmatrix} \bar{b}' \otimes \beta + I \otimes \beta\bar{b} & I \otimes \beta \\ \bar{h}' \otimes \beta & I \otimes \beta\bar{b} \end{bmatrix}, \tag{114}$$

$$DT_f(\bar{b}) = I \otimes \beta\bar{b}, \tag{115}$$

$$DT_e(\bar{b}) = \beta\bar{b}. \tag{116}$$

We conjecture that the analogous proposition to (2.2.8) holds, i.e. a necessary condition for E-stability of the set of general form representations $\Xi_{(\bar{b},\bar{h})}$ is that none of the eigenvalues of $DT_a(\bar{b})$, $DT_{(b,h)}(\bar{b},\bar{h})$, $DT_c(\bar{b})$ and $DT_e(\bar{b})$ has a real part larger than 1. We presume that this condition is also sufficient if apart from m eigenvalues of $DT_e(\bar{b})$ equal to 1 the eigenvalues of $DT_a(\bar{b})$, $DT_{(b,h)}(\bar{b},\bar{h})$, $DT_c(\bar{b})$ and $DT_e(\bar{b})$ have real parts different from 1, where $m = \dim(\{x|(DT_e(\bar{b}) - I)x = 0\})$. Currently our results are limited to purely forward looking models. In case of a purely forward looking model the T-map simplifies to (99), (100), (101) and (106). As in case of common factor representations let $\Xi_{\bar{b}} = \{\xi | T(\xi) = \xi \wedge b = \bar{b}\}$ be the set of REERs corresponding to one particular solution to (100). Now the following proposition is an immediate consequence of proposition 2.2.8.

PROPOSITION 2.2.10. *Assume that the information set at time n is* $I_n = \{1, y_{n-1}, w_n, \varepsilon_n\}$ *and that the model is purely forward looking* ($\delta = 0$). *A necessary condition for E-stability of a set of general form representations* $\Xi_{\bar{b}}$ *is that none of the eigenvalues of* $DT_a(\bar{b})$, $DT_b(\bar{b})$, $DT_c(\bar{b})$ *and* $DT_e(\bar{b})$ *has a real part larger than 1.*

PROOF. Immediate from proposition 2.2.8. □

Note that the previous proposition does not give sufficient conditions for E-stability. The following proposition shows that general form representations cannot be E-stable in a purely forward looking model.

PROPOSITION 2.2.11. *Assume that the information set at time n is* $I_n = \{1, y_{n-1}, w_n, \varepsilon_n\}$ *and that the model is purely forward looking* ($\delta = 0$). *Then all general form representations with* $\bar{b} \neq 0$ *are E-unstable.*

PROOF. See appendix A.2. □

The qualification $\bar{b} \neq 0$ is necessary, because there is a unique general form representation with $\bar{b} = 0$, which is identical to the MSV representation. As noted above the MSV representations are nested in the functional form of both common factor and general form representations. However, all proper sunspot representations (i.e. those with $\bar{b} \neq 0$) of the general form are always E-unstable. We conjecture that this instability result also holds for all models with a sufficiently low degree of inertia.[10] Corollary 2.2.9 and proposition 2.2.11 show that common factor representations are E-stable whenever the MSV representation is E-stable whereas general form representations are always E-unstable. As we have argued above both results extend to models with a sufficiently low degree of inertia. This finding rationalises why common factor representations have often been found to be E-stable in the literature in contrast to their counterparts of the general form.

2.3. Convergence of RLS

The previous section has focused exclusively on the stability of the differential equation (109). Due to its importance for the convergence of RLS it has been labelled E-stability (cf. definition 2.2.7). The ODE (109) has been shown to be a good approximation for the average asymptotic behaviour of least squares learning in case of regular REE (cf. e.g. Evans and Honkapohja, 2001). The main technical difficulty in proving convergence in the presence of sunspots is that there is an unbounded continuum of equilibria. All known convergence results require boundedness of the sequence of estimates. We will assume that the estimates of the agents are projected on some compact set, which contains a non-empty set of sunspot equilibria and ask wether convergence occurs to some equilibrium within this set. Of course such a procedure does exclude equilibria that are far away from the initial beliefs of the agents. However our simulations in section 2.4 indicate that convergence generally occurs to an equilibrium point close to the initial beliefs of the agents even without imposition of a projection facility.

The second part of the problem is that there is a continuum of equilibria instead of an isolated equilibrium. We apply the results of Delyon (1996), which extend earlier results of Metivier and Priouret (1984) and Benveniste et al. (1990) to cover continua of equilibria.

2.3.1. A Convergence Theorem for Recursive Stochastic Algorithms. We consider recursive stochastic algorithms of the form

$$\theta_n = \pi\Big(\theta_{n-1} + \gamma_n \psi_n\big(H(\theta_{n-1}, X_n) + \gamma_n \rho_n(\theta_{n-1}, X_n)\big)\Big), \tag{117}$$

where $\theta_n \in \mathbb{R}^k$ is a vector of parameters, $X_n \in \mathbb{R}^l$ is a vector of state variables, γ_n is a scalar sequence, $H : \mathbb{R}^k \times \mathbb{R}^l \to \mathbb{R}^k$ and $\rho_n : \mathbb{R}^k \times \mathbb{R}^l \to \mathbb{R}^k$ are both functions of the parameters and the states and π and ψ_n are projections. The projection π is in general necessary to assure that the parameters θ_n remain in some compact set $Q \subset \mathbb{R}^k$, whereas the projections

[10]This remark is based on the assumption that an unstable root in (114) leads to E-instability of general form representations in models with inertia. However, as we have remarked above it remains to be proven that this condition is necessary for E-stability of general form representations.

ψ_n guarantee that $|\theta_n - \theta_{n-1}| \to 0$. We need the following assumptions on the stochastic algorithm (117)

(A1) $\gamma_n \geq 0, \sum \gamma_n = \infty, \sum \gamma_n^2 < \infty$, and $\sum |\gamma_n - \gamma_{n+1}| < \infty$,

(A2) there exist C_1, C_2, p_1, p_2 such that for all $\theta \in Q$, $x \in \mathbb{R}^l$

$$\begin{aligned} |H(\theta,x)| &\leq C_1(1+|x|^{p_1}), \\ |\rho_n(\theta,x)| &\leq C_2(1+|x|^{p_2}), \end{aligned}$$

(A3) $H(\theta,x)$ is twice continuously differentiable with bounded second derivatives on Q,

(A4) for some $C > 0$ and $\eta < 1$ let $B_n = \{\theta \in \mathbb{R}^k | |\theta| < C\gamma_n^{-\eta}\}$ then for all n, θ: $\psi_n(\theta) \in B_n$ and $\psi_n(\theta) = \theta$ if $\theta \in B_n$.

Assumption (A1) guarantees that the gain sequence decreases fast enough to obtain convergence but not too fast, so that convergence to a non-equilibrium point can be excluded. It is readily satisfied for $\gamma_n = 1/n$. (A2) and (A3) are regularity assumptions on the functions H and ρ_n. As mentioned above the projections ψ_n guarantee that $|\theta_n - \theta_{n-1}| \to 0$. We assume that the dynamics of the state vector follows a conditionally linear process

$$X_n = A(\theta_{n-1})X_{n-1} + B(\theta_{n-1})W_n, \tag{118}$$

which satisfies the set of assumptions (B).

(B1) W_n is iid with finite absolute moments,

(B2) there exist constants K_1, K_2, K_3 and $0 < \omega < 1$ such that for all $\theta \in Q$, all $n \geq 0$

$$|A(\theta)| \leq K_1 \quad |A(\theta)^n| < K_2\omega^n, \quad |B(\theta)| \leq K_3$$

for some matrix norm $|\cdot|$, and $A(\theta)$ and $B(\theta)$ satisfy Lipschitz conditions on Q with the constants K_4 and K_5.

Assumption (B2) is satisfied if the system matrices are continuously differentiable functions of θ and the state process is asymptotically stationary. Under assumptions (B1) and (B2), the state vector X_n converges in L_q for all $q > 0$ to the random variable [11]

$$X_\infty(\theta) = \sum_{k=1}^{\infty} A^k(\theta)B(\theta)W_k.$$

The central idea behind the convergence of the recursive stochastic algorithm can be seen most easily ignoring the projections and rewriting (117) as [12]

$$\frac{\theta_n - \theta_{n-1}}{\gamma_n} = H(\theta_{n-1}, X_n) + \gamma_n \rho_n(\theta_{n-1}, X_n) \tag{119}$$

[11] For a proof cf. Evans and Honkapohja (2001) p.129.

[12] Examples for recursive stochastic algorithms are recursive least squares and stochastic gradient learning.

For n large enough ρ_n is negligible. As γ_n goes to zero the left hand side is well approximated by the differential of θ, whereas the right hand side can be approximated by $h(\theta) = E[H(\theta, X_\infty(\theta))]$, where θ is kept fixed and the expectation is taken with respect to the distribution of $X_\infty(\theta)$ given by (2.3.1). This reasoning leads to the interest in the study of the mean ODE or mean dynamics, which is the differential equation

$$\dot{\theta} = h(\theta). \tag{120}$$

As shown by Evans and Honkapohja (1998) $h(\theta)$ is well defined and locally Lipschitz under assumptions (A1-2) and (B1-2). The name mean dynamics derives from the fact that it describes the average asymptotic behaviour of the recursive stochastic algorithm (117). As we will see below the recursive stochastic algorithm converges locally to a set of equilibrium points of the mean ODE provided the set is locally asymptotically stable. Before we present the convergence proof we need two more assumptions.

(Proj) Q is a bounded set, π is a Lipschitz continuous function such that $\pi(\theta) \in Q$ for all θ and $\pi(\theta) = \theta$ if $\theta \in Q$. There exists a neighbourhood O of Q such that for some $\delta > 0$

$$\langle \nabla V(\theta), \pi(\theta) - \theta \rangle < -\delta |\pi(\theta) - \theta|, \quad \text{for } x \in O \backslash Q, \tag{121}$$

where V is the Lyapunov function in (Stab),

(Stab) there exists a non-negative C^1 function V and a set of stationary points $\Theta \subset O$ closed with respect to O (the same as in (Proj)) such that

- h is continuous, and $\langle \nabla V(\theta), h(\theta) \rangle < 0$ if $\theta \in O \backslash \Theta$,
- $V(\theta)$ is constant on any connected subset of Θ.

Assumption (Proj) guarantees that θ_n remains within a compact set Q and that there are no additional rest points of the mean dynamics at the boundary of Q that are introduced by the projection (condition 121). The stability condition (Stab) postulates that Θ is a locally asymptotically stable set of equilibrium points of the mean dynamics.[13] We are now able to prove the following convergence result.

PROPOSITION 2.3.1. *Let θ_n be given by algorithm* (117). *Assume that the state X_n follows the conditionally linear process* (118) *and that (A), (B), (Proj) and (Stab) hold. Then θ_n converges a.s. to Θ. Furthermore X_n is bounded in L_q for all $q > 0$ and the projections ψ_n are made only a finite number of times.*

PROOF. See appendix A.2. □

[13] Wilson (1969) and Lin et al. (1996) prove the existence of such a Lyapunov function V provided the set Θ is asymptotically stable.

REMARK 2.3.2. The projections ψ_n are not necessary if the state dynamics (118) is independent of θ.

Probably the projections ψ_n are not necessary at all, however the proof without these projections requires significantly more technical details (cf. Benveniste et al., 1990, chapter 3). The assumption of the projection facility π (Proj) is wide spread in the literature (cf. e.g. Marcet and Sargent, 1989b). It has been criticised because it is hardly verifiable in applications.[14] As a consequence convergence results that do not require the projection π have been developed (cf. e.g. Evans and Honkapohja, 1998). Without assumption (Proj) it is generally not possible to obtain convergence with probability one. Yet, for large n, the gain γ_n becomes very small and the probability of convergence approaches one. We conjecture that it is also possible to derive such a result for the convergence to sunspot equilibria, but this is beyond the scope of this thesis.

2.3.2. Application to Least Squares Learning. We will now apply proposition 2.3.1 to least squares learning of common factor representations. The application to general form representations is completely analogous. Under real time learning the agents update their beliefs about the coefficients ξ_n of the perceived law of motion

$$y_n = \xi_n' z_n \tag{122}$$

every period given the latest available information $I_n = \{y_{k-1}, w_k, \varepsilon_k; k \le n\}$. Recall that for a common factor representation $\xi_n' = (a_n, b_n, c_n, d_n)$ and $z_n' = (1, y_{n-1}', w_n', \zeta_n')$. Given the current beliefs ξ_n of the agents the actual law of motion of the economy evolves according to

$$y_n = T(\xi_n)' z_n, \tag{123}$$

where $T(\xi) = (T_a(a,b), T_b(b), T_c(b,c), T_d(b,d))$ and the mappings $T_v(\cdot)$ are given by (99)–(102). We assume that the learning algorithm of the agents is given by the following recursive stochastic algorithm

$$\theta_n = \pi\left(\theta_{n-1} + \frac{1}{n}\psi_n\begin{pmatrix}\operatorname{vec}(R_{n-1}^{-1} z_{n-1} z_{n-1}'(T(\xi_{n-1}) - \xi_{n-1}))\\ n/(n+1)\operatorname{vec}(z_n z_n' - R_{n-1})\end{pmatrix}\right), \tag{124}$$

where $\theta_n = \operatorname{vec}(\xi_n, R_n)$ and the projections π and ψ_n satisfy assumptions (Proj) and (A4) respectively. This learning algorithm is a least squares algorithm augmented with two projections π and ψ_n. If $\pi(\theta) = \theta$ and $\psi_n(\theta) = \theta$ for all n, θ, then (124) is identical to

[14] In general it is not possible to obtain the explicit functional form of the Lyapunov function V and in these cases it is not possible to verify that condition (121) holds for a given set Q. If condition (121) fails and the remaining conditions in proposition 2.3.1 are all satisfied then θ_n converges either to Θ or to the boundary of the set Q.

recursive least squares and given appropriate initial conditions, ξ_n equals the OLS estimator while R_n is the corresponding moment matrix[15]

$$\xi_n = \left(\sum_{i=0}^{n-1} z_i z_i'\right)^{-1} \sum_{i=0}^{n-1} z_i y_i' \quad \text{and} \quad R_n = \frac{1}{n}\sum_{i=0}^{n} z_i z_i'. \tag{125}$$

In general our convergence proof will require that the projections are not simply identities.[16] Yet, as we will see in section 2.4, our simulation results suggest that in the New Keynesian model convergence obtains even without imposition of a projection facility for initial conditions in a neighbourhood of the set of common factor representations. The learning algorithm (124) is a function of the state variables $X_n' = (1, y_{n-1}', w_n', \zeta_n', y_{n-2}', w_{n-1}', \zeta_{n-1}')$, which follow the conditionally linear Markov process

$$X_n = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ T_a(a_{n-1}, b_{n-1}) & T_b(b_{n-1}) & T_c(b_{n-1}, c_{n-1}) & T_d(b_{n-1}, d_{n-1}) & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varphi & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \end{bmatrix} X_{n-1} + \begin{bmatrix} 1 \\ 0 \\ \vartheta_n \\ \varepsilon_n \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{126}$$

Conditional on ξ_{n-1} the process (126) is a Markov representation of the actual law of motion of the economy under learning. Together equations (124) and (126) give a complete characterisation of the model (93)–(94) under learning. Before we state the convergence result define $\Theta_{\bar{b}} = \{\theta = \text{vec}(\xi, R) | \xi \in \Xi_{\bar{b}} \wedge R = \bar{M}(\xi)\}$, where $\bar{M}(\xi) = E[z_n(\xi) z_n(\xi)']$.[17] Recall that $\Xi_{\bar{b}}$ is the set of common factor representations given one particular $\bar{b}$ (cf. section 2.2). Consider a $\theta = \text{vec}(\xi, R) \in \Theta_{\bar{b}}$. The ξ-component of θ corresponds to a particular common factor representation, while the R-component is the expected value of the moment matrix of the regressors z_n conditional on ξ. Now we can apply proposition 2.3.1 to prove the convergence of least squares learning.

PROPOSITION 2.3.3. *Consider the learning algorithm* (124) *and the state dynamics* (126). *Assume that the projections* π *and* ψ_n *satisfy assumptions (Proj) and (A4) respectively. Assume that* $\Theta_{\bar{b}}$ *is a locally asymptotically stable set of the mean ODE* (120). *Let* Q *(the same as in (Proj)) be a compact set such that*

(1) $Q \cap \Theta_{\bar{b}} \neq \emptyset$,

[15]Note that ξ_n includes only information up to period $n-1$, because y_n is not included in the information set I_n (for a discussion of this assumption cf. section 2.2).

[16]The projections are discussed in detail in the previous paragraph.

[17]The expectation is taken with respect to the unique stationary measure of the Markov process (126) holding ξ constant.

(2) *the inverse of R is well defined,*

(3) *the roots of $T_b(b)$ are inside the unit circle,*

(4) *Q is a subset of the domain of attraction of $\Theta_{\bar{b}}$,*

then θ_n converges to $\Theta_{\bar{b}}$ a.s. and the projections ψ_n are made only a finite number of times.

PROOF. See appendix A.2. □

Proposition 2.3.3 is the main proposition of this section. It states that the learning algorithm converges almost surely to the set $\Theta_{\bar{b}}$ if it is locally asymptotically stable with respect to the mean ODE $\dot{\theta} = h(\theta)$. In other words ξ_n converges to a common factor representation $\xi \in \Xi_{\bar{b}}$ and R_n converges to the corresponding moment matrix $\bar{M}(\xi)$. [18] However, convergence with probability one requires a number of assumptions on the set Q. The first one is obvious. Convergence to the set $\Theta_{\bar{b}}$ can of course occur only if the projection set Q includes a non-empty subset of equilibrium representations. It is also clear that the domain of the moment matrix R must be chosen such that the inverse is well defined. The third assumption requires that $\Xi_{\bar{b}}$ is a set of stationary common factor representations. It is readily satisfied in a neighbourhood of $\Theta_{\bar{b}}$ if all eigenvalues of $\bar{b}$ are inside the unit circle. The fourth assumption assures that the learning algorithm stays within the domain of attraction of the mean dynamics. It holds in a neighbourhood of $\Theta_{\bar{b}}$, if $\Theta_{\bar{b}}$ is an asymptotically stable set of the mean ODE.

As shown in proposition 2.3.3 convergence hinges on the question wether the set $\Theta_{\bar{b}}$ is a locally asymptotically stable set of the mean ODE (120). The mean ODE of the learning algorithm (124) can be computed as

$$\dot{\xi} = R^{-1}\bar{M}(\xi)(T(\xi)-\xi) \tag{127}$$

$$\dot{R} = \bar{M}(\xi)-R. \tag{128}$$

As noted by Marcet and Sargent (1989b), given a vector of beliefs ξ, the second component of the mean dynamics (128) is globally asymptotically stable, so that $R \to \bar{M}(\xi)$. Setting $R = \bar{M}(\xi)$ in the first component of the mean dynamics (127) yields the smaller ODE (109), which defines E-stability (cmp. definition 2.2.7). For regular equilibria it has been established formally by Marcet and Sargent (1989b) that E-stability is equivalent to stability of the mean ODE. For the same technical reasons which make the analysis of E-stability extremely complicated in case of sunspot equilibria it is hard to establish a connection between E-stability and stability of the mean ODE. We conjecture that E-stability and stability of the mean ODE are still equivalent, but in the presence of sunspots this remains to be proven. Section 2.4 applies our results to a prime example of the literature

[18] We conjecture that it is also possible to derive an instability theorem, which states that there is convergence with probability zero to a non-equilibrium point. We leave this for future research.

on sunspots in monetary policy and illustrates numerically the convergence of RLS to a common factor sunspot representation under E-stability.

2.4. Application

We will now apply the results of the previous sections to study the convergence of least squares learning in a standard New Keynesian model.

$$z_n = E_n z_{n+1} - \sigma^{-1}(R_n - E_n \pi_{n+1}) + g_n, \tag{129}$$

$$\pi_n = \kappa z_n + \beta E_n \pi_{n+1} + u_n. \tag{130}$$

Here z_n is the output gap, π_n is the inflation rate and R_n is the nominal interest rate. The demand shock g_n and the cost push shock u_n are assumed to follow AR(1) processes with damping parameters $0 \leq \rho_g \leq 1$ and $0 \leq \rho_u \leq 1$. The model (129)–(130) consists of a forward looking dynamic IS-curve and a Philips curve. Monetary policy is assumed to set interest rates according to the forward looking rule

$$R_n = \phi_\pi E_n \pi_{n+1} + \phi_y E_n y_{n+1}. \tag{131}$$

It is a well known fact that forward looking interest rate rules are susceptible to sunspots. As we have shown in section 2.2 general form representations are always E-unstable in purely forward looking models, whereas common factor representations are E-stable whenever the MSV representation is E-stable. In contrast to the previous literature we have also shown that the well known conditions on the derivatives of the T-map are necessary and sufficient for E-stability (proposition 2.2.8). The following proposition applies these results to the model (129)–(131).

PROPOSITION 2.4.1. *Under a forward looking interest rate rule of the form* (131) *there are sunspot equilibria with E-stable common factor representations if and only if*

$$\phi_\pi + (1-\beta)\kappa^{-1}\phi_y > 1, \tag{132}$$

$$\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_y > 2\sigma_y. \tag{133}$$

General form representations with $b \neq 0$ *are always E-unstable.*

PROOF. Proof see appendix $A.2$. □

Proposition 2.4.1 shows that common factor representations are E-stable provided the Taylor principle holds (132) and the nominal interest rate responds sufficiently strong to changes in the output gap (133). Moreover all general form representations, which do not coincide with the MSV representation (i.e. with $b \neq 0$) are E-unstable. We conjecture that E-stability of a set of sunspot representations in the sense of definition 2.2.7 is equivalent to local stability under least squares learning. From proposition 2.3.3 we know that least squares learning converges to a sunspot equilibrium provided the mean ODE (120) is asymptotically stable. The essential step which is still missing to proof the conjecture is the proof that E-stability is equivalent to local asymptotic stability of the mean ODE. We will now conduct a small simulation study for the New Keynesian model laid out above. The results seem to support our hypothesis that E-stability of a set of common factor representations is both necessary and sufficient for the local convergence of least squares learning towards the set.[19]

We calibrate the model with standard values from the literature. Qualitatively the results do not seem to be sensitive to the exact values which are chosen. We set the discount factor to $\beta = 0.99$ and use the estimates $\kappa = 0.024$ and $\sigma = 0.157$ of Rotemberg and Woodford (1998) and Rotemberg and Woodford (1999). The damping parameters of the demand shock ρ_g and the cost push shock ρ_u are set to 0.95 and 0.9. The standard deviations of the shocks are both set to 0.05. The coefficients of the monetary policy rule are calibrated to $\phi_\pi = 1.5$ and $\phi_y = 0.5$. By proposition 2.4.1 this calibration yields an E-stable set of common factor representations, so we would expect convergence of recursive least squares towards some common factor representation within this set. We initialise the sunspot multiplier at $d = (1.5, 0.2)'$. All remaining estimates are initialised at some values close to the common factor representation given by proposition 2.2.6. Qualitatively the results do not seem to depend on the starting values. Figure 1 plots the sequence of recursive least squares estimates for d obtained from a simulation of the model economy over 5000 periods.[20] The upper two graphs show the evolution of the impact multipliers of the sunspot shock on the output gap (on the left) and on inflation (on the right). It can be seen that the impact of the sunspot on the output gap is much larger than on inflation. Moreover both coefficients seem to converge. The lower two figures contain a scatter plot of the two multipliers against each other. The straight line is the continuum of sunspot equilibria. The left panel plots the whole sequence of estimates, whereas the right panel plots only the last 1000 estimates. In the right panel it can be seen that the estimates seem to settle down to a single point on the line. Starting from some belief close to the set of sunspot equilibria, there seems to be convergence to a particular sunspot equilibrium in the neighbourhood of the initial belief. The point of convergence depends on the initial value for d and on the realised sequence of shocks. Remarkably, in all simulations we have conducted convergence obtained without imposition of a projection facility.

[19] One regularity condition which has to hold for this to be true is that all common factor representations within the set are asymptotically stationary.

[20] We only plot the evolution of the estimates for d because all remaining parameters converge fairly quickly to their equilibrium values.

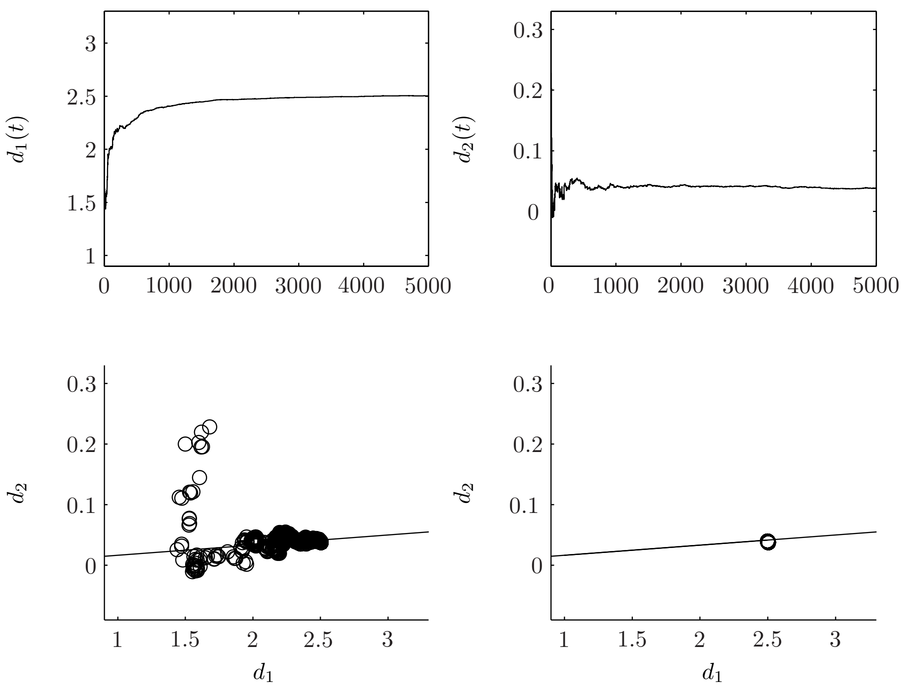

FIGURE 1. *recursive least squares estimates of the sunspot multiplier d*

If the agents stipulate that the equilibrium may change over time, then they may want to keep track of these changes using a constant gain algorithm. Figure 2 plots the sequence of estimates in case the agents use a constant gain algorithm to estimate both the common factor representation and the resonant frequency of the sunspot.[21] Interestingly the estimates initially tend towards the set of equilibria. In contrast to the case of recursive least squares, the beliefs do not settle down on a single value, but wander around within the set of equilibrium values as can be seen from the lower right panel of figure 2.[22] Again this panel plots only the last 1000 estimates. The upper two panels of the figure show that the impact of the sunspot fluctuates considerably under constant gain learning. Recall that the straight line which represents the set of all equilibrium values of the sunspot multiplier d contains the origin $d = 0$. Therefore, under adaptive learning, there may be periods when sunspots have large effects on the economy and other periods when they are negligible. As a consequence there may be periods with very high and periods with low output gap

[21]The resonant frequency can be estimated by a regression of ξ_n on ξ_{n-1}.

[22]Note that the effect, that the estimates wander around within the set of REERs disappears if the resonant frequency is fixed at the true value. In this sense it is not innocuous to assume that the resonant frequency is known as it is often done in the learning literature.

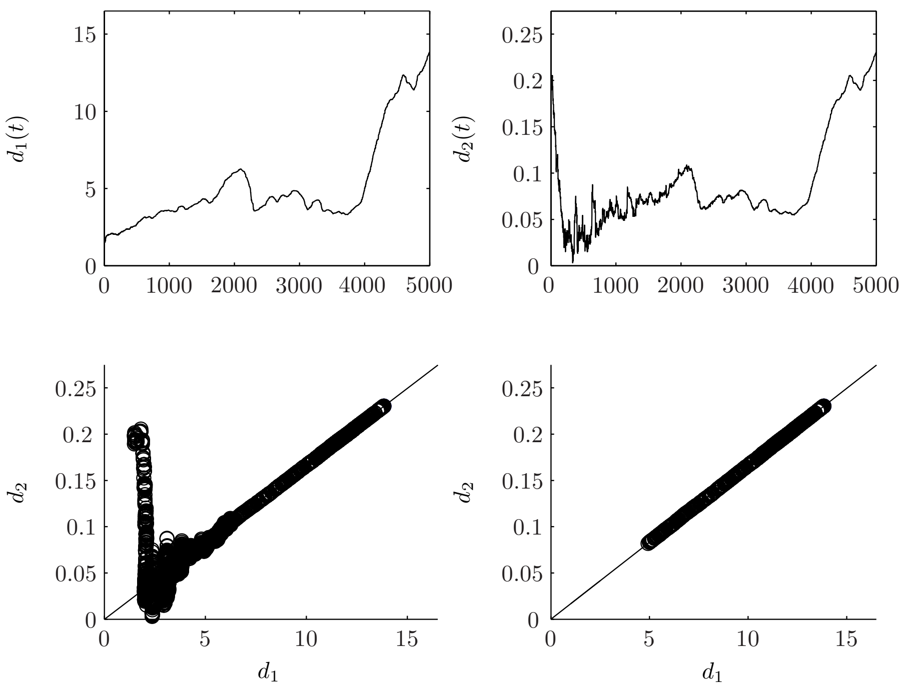

FIGURE 2. *recursive estimates of the sunspot multiplier d under constant gain learning*

volatility.[23] Finally we want to remark that we have conducted numerous simulations for choices of policy parameters, which yield E-unstable common factor or general form representations. [24] In all these simulations least squares learning has diverged fairly quickly suggesting that E-instability implies that least squares learning converges with probability zero.

2.5. Conclusions

We have provided several new insights on E-stability and the convergence of least squares learning to sunspot equilibria. In section 2.2 we have established necessary and sufficient conditions for E-stability of common factor representations. In the presence of sunspots standard stability results for differential equations are not applicable, which is why a formal derivation of these conditions has been missing in the literature so far. In a purely forward looking model common factor representations are E-stable if and only if the

[23]Of course the same is true for inflation, but – at least in our calibration – the impact of the sunspot on the output gap is much larger.

[24]Recall that general form representations are E-unstable for all choices of the policy parameters (cf. proposition 2.4.1).

equilibrium is indeterminate and the MSV representation is E-stable. We have also proven an instability result, which demonstrates that general form representations are always E-unstable in purely forward looking models. We argue that both results extend to any model with a sufficiently low degree of inertia. This explains why common factor representations are often E-stable while general form representations are E-unstable. In section 2.3 we have shown that under standard regularity assumptions least squares learning converges to rational expectations provided the so called mean dynamics is stable. We use an approach by Delyon (1996) to show almost sure convergence of least squares learning (augmented with projection facilities) to a set of sunspot equilibria provided the set is locally asymptotically stable under the mean ODE. For locally isolated equilibrium representations (e.g. any representation of a locally determinate equilibrium) it is well known in the literature that stability of the mean ODE is equivalent to E-stability. However, sunspot representations are in general not locally isolated. Instead there is a continuum of equilibrium representations. We conjecture that stability of the mean ODE and E-stability are still equivalent, but this remains to be proven.

Section 2.4 applies our E-stability results to a New Keynesian model with a forward looking interest rate rule, which is the prime example of sunspots in monetary policy. From section 2.4 we know that there is an E-stable set of common factor representations provided the interest rate rule satisfies the Taylor principle and the nominal interest rate is very sensitive to changes in the output gap. Moreover general form representations are never E-stable. In order to assess our conjecture that E-stability is a necessary and sufficient condition for the convergence of least squares learning we conduct a small simulation study. It turns out that least squares seems to converge the set of common factor representations provided the set is E-stable. Convergence takes place for all starting values in a neighbourhood of the set even without imposition of a projection facility. The particular limit point within the set depends on the initial beliefs of the agents and the realised shock sequence. In addition least squares seems to diverge if the agents try to learn an E-unstable general form or common factor representation. If agents try to track parameter changes using a constant gain learning algorithm the beliefs of the agents wander around within the set of common factor representations. In this case learning can explain that there are periods where the effect of sunspots on economic variables is large and other periods where it is negligible. As a consequence there may be periods with high and periods with low volatility in the data. We conjecture that this is a more general phenomenon, which has not been observed previously and which is likely to occur in other models too.

CHAPTER 3

Sticky Information, Adaptive Learning and the Great Moderation

3.1. Introduction

The substantial increase in macroeconomic performance across many industrialised countries over the last decades has triggered a lot of work on the underlying causes. It is widely accepted that the fall in average inflation and inflation volatility is mostly due to improvements in the conduct of monetary policy. Most economists attribute the rise in average inflation in the 1960s and 1970s to the attempt of monetary policy to exploit a long run trade-off between inflation and the output gap or unemployment (cf. Taylor, 1996, proposition 1). In the United States, monetary policy making seems to have changed significantly with the appointment of Paul Volcker as chairman of the Federal Reserve Bank. Clarida et al. (2000) claim in this regard that 'not until Volcker took office did controlling inflation become the organising focus of monetary policy'. Since then it is widely believed that in the long run monetary policy determines the average rate of inflation and has no real effects, while there is a short run trade-off between the variability of inflation and the output gap or unemployment. Thus there is a trade-off in the volatilities instead of the levels of inflation and aggregate economic activity (cf. Taylor, 1996, proposition 2). Despite the considerable agreement that improvements in the conduct of monetary policy led to the stabilisation of inflation, the simultaneous decline in the volatility of aggregate economic activity is very controversial.

One widespread view holds that monetary policy has been too accommodative with respect to inflation expectations in the pre-Volcker period increasing interest rates less than one-for-one with inflation expectations. Such a regime may lead to self-fulfilling expectations because an increase in inflation expectations decreases the real interest rate, which in turn increases economic activity, thus justifying the increased inflation expectations. In contrast, with the beginning of the Volcker period the Federal Reserve managed to reduce inflation expectations and bring down average inflation as well as inflation volatility by increasing rates more than one-for-one with inflation. As argued by Clarida et al. (2000) this fundamental change in monetary policy has led to greater macroeconomic stability. Apart from the possibility of self-fulfilling expectations some economists attribute the moderation of the volatility of the output gap to the decline in the average inflation rate. They argue that low inflation reduces nominal distortions arising from taxation and insofar as low inflation means low expected inflation, it increases the possibilities of the Fed to react to unforeseen events (cf. Summers, 2005). In contrast McConnell and Perez-Quiros (2000) claim that the reduction in the volatility of economic activity is completely unrelated to monetary policy. They blame better practices in inventory management that have come along with the increased adoption of information technology since the 1980s to be the cause for the increased stability. Finally the great macroeconomic instability during the 1970s and 1980s may simply have been bad luck. Among others Gordon (2005) claims that a series of large supply shocks are responsible for the bad performance in that period.

This chapter focuses on the interaction of monetary policy and sticky information in the presence of adaptive learning. The hypothesis we investigate is whether the break in monetary policy with the appointment of Paul Volcker may have contributed to the Great Moderation. The basic effect has been discovered by Branch et al. (2004). In a model with a trade-off between inflation and output volatility an increased desire to stabilise inflation leads to a direct decrease in inflation and increase in output volatility. In the presence of endogenous inattention there is however also an indirect effect, which reduces both volatilities. Depending on the size of both effects the volatility trade-off may breakdown.

The model of Branch et al. (2004) builds on the sticky information Philips curve suggested by (Mankiw and Reis, 2002). In contrast to the sticky price literature, in the sticky information approach agents set their prices optimally at each point in time. However, they are assumed to update their information sets only sporadically, so that their decisions may be based on outdated information. This assumption is usually justified by the presence of limited capabilities and resources to process information as well as costs to collect information. In Branch et al. (2004) suppliers endogenously choose the probability at which they update their information sets given that updating information is costly. In such an environment an increase in inflation hawkishness on part of the central bank can lead to a simultaneous decline of inflation and output volatility. The immediate consequence is a decline in the volatility of inflation and a rise in the volatility of the output gap. The decline of inflation volatility in turn lowers the benefits of suppliers to update their information sets. Given that the costs for information acquisition remain unchanged they will endogenously choose to update their information less frequently, which amplifies the decrease in the volatility of inflation and may even lead to an overall decline in the volatility of the output gap. Branch et al. (2006) show that their model is stable under adaptive learning. However, they also emphasise that their model is too simple to provide a description of the actual historical experience and so they do not attempt to calibrate the model. The model of Branch et al. (2006) is a version of the Mankiw and Reis (2002) model, which is not microfounded. In order to clearly separate and illustrate the effects of sticky information the model consists of a version of a sticky information Philips curve, a quantity equation and a simple money supply rule.

In this chapter we build on the subsequent work of Reis (2006), Mankiw and Reis (2006a) and Mankiw and Reis (2006b) who extend the sticky information approach to a general equilibrium framework. In contrast to Mankiw and Reis (2006a) and Mankiw and Reis (2006b) we assume only suppliers choose their level of inattention, while consumers and workers are completely attentive. [1] The model consists of a sticky information Philips curve, an intertemporal IS curve and an interest rate rule. We find that for plausible parameter values the model is capable to generate a decline in both volatilities, which is comparable to that found in the data. The fact that the agents need to learn about changes in monetary policy means that their attentiveness adapts only gradually to the new equilibrium. This introduces a significant short run trade-off between the volatility of inflation and the output gap. However, in the long run the model seems to converge to the new rational expectations equilibrium. As in the data we find that the volatilities of inflation and the output gap are highly correlated. Moreover the correlation between the volatilities is much larger than the correlation between the levels. The following section 3.2 discusses some empirical observations regarding the Great Moderation. Section 3.3 derives the model and section 3.4 describes the rational expectations equilibrium and endogenous inattention. Section 3.5 introduces adaptive learning. The simulation results of the calibrated economy can be found in section 3.6, while section 3.7 concludes.

3.2. Empirical Facts

The sharp decline in US business cycle volatility is generally dated to the mid 1980s. Some economists argue that the decline occurred rather suddenly, presumably in the first quarter of 1984 (cf. McConnell and Perez-Quiros (2000) or Kim and Nelson (1999)). Others like Blanchard and Simon (2001) claim that there has been a steady decline over several decades. Although the timing differs from country to country, most industrialised countries have experienced a similar decline in the volatility of their business cycles during the last decades (cf. Summers, 2005).

Figure 1 shows 230 quarters of data on US inflation, output growth and the output gap from 1949:Q1 to 2006:Q2. Inflation is computed as the year over year percentage change in the GDP price deflator. Output growth is the percentage year over year increase in real GDP. Real GDP and the price deflator are taken from the FRED. The output gap is the percentage deviation of real GDP from its potential, where potential real GDP is taken from the CBO. It is apparent from figure 1 that there is a decline in the level of inflation

[1] The assumption of completely attentive consumers and producers reduces the dimension of the problem considerably, especially if agents are learning and choose their updating probability endogenously.

in the mid 1980s, which is generally associated with a break in monetary policy under Fed chairman Paul Volcker. Before 1984 the inflation rate moves in a range between -2 and 10 percent with a mean of 3.8 percent compared to a range of 1 to 4 percent and a mean of 2.5 percent after 1984. Along with the decline of the level and the volatility of inflation, output growth has become much more stable. Booms tend to last longer, while there is a narrower gap between growth rates in booms and recessions after 1984 (Kim and Nelson, 1999). The behaviour of the output gap confirms the increased stability of aggregate economic activity following the mid 1980s.

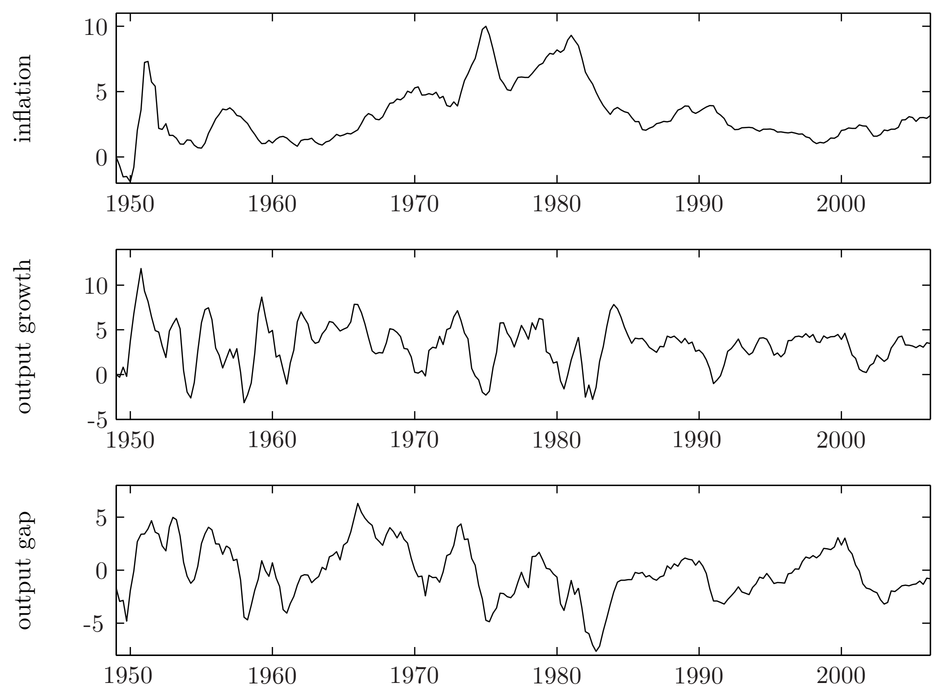

FIGURE 1. *US data from 1949:Q1 to 2006:Q2*

Figure 2 presents the 20-quarter rolling standard deviations of inflation, output growth and the output gap. It confirms the impression from figure 1 that there is a sharp decline in the volatilities of all three quantities. Compared to figure 1 the decline is delayed by a few years and occurs in the late 1980s. The reason is that the rolling window is a backward looking of the standard deviation. If we split the sample in the first quarter of 1984 the results are even more dramatic. The standard deviation of inflation before 1984 is 2.67 compared to 0.76 afterwards. Thus it has dropped roughly to a quarter. The standard deviations of output growth and the output gap both have dropped by roughly one half from 2.94 and 2.99 to 1.52 and 1.51 respectively.

Figures 1 and 2 suggest the decline in the level of inflation and inflation volatility may have contributed to the moderation of business cycle volatility. The correlation between the level of inflation and the volatility of output growth or the output gap is 0.13 and 0.18 respectively. Thus, the co-movement of the level of inflation and the volatility of economic activity is positive but not particulary strong. Instead the correlation between the standard deviations of inflation and output growth is 0.65 – between inflation and the output gap it is 0.77. So the co-movement of the volatilities reinforces the impression from figure 2 that the Great Moderation of economic activity is linked to the moderation of inflation. It is widely recognised that a break in the Fed's conduct of monetary policy under chairman Paul Volcker helped to bring down both average inflation and inflation volatility. Together with the estimate that the break from high to low business cycle volatility occurred most probably in the first quarter 1984, this is strong evidence that the change in monetary policy is also part of the story of the Great Moderation. The model we will set up in the next section establishes a strong link between the volatility of inflation and the volatility of economic activity.

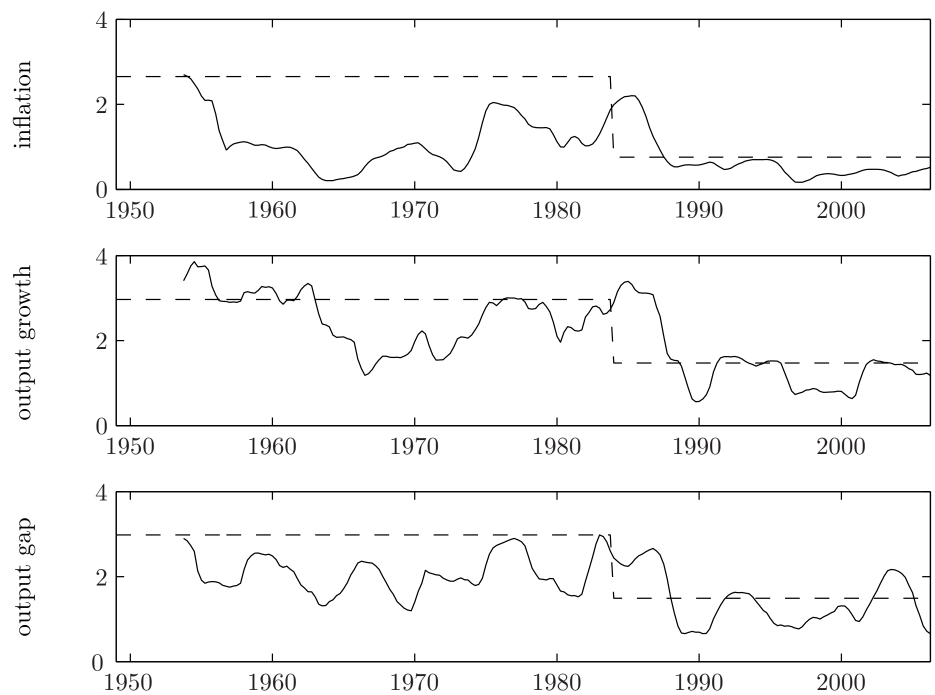

FIGURE 2. *standard deviations before/after 1984:Q1 vs. 20-quarter rolling standard deviations*

3.3. The Model

In this section we build a simple model for the inflation rate and the output gap. The supply side of the model evolves according to a sticky information Philips curve as introduced in Mankiw and Reis (2002). The demand side on the other hand follows a standard intertemporal IS curve. The basic model can be seen as a special case of Mankiw and Reis (2006b), where consumers and workers are assumed to be completely attentive. The advantage of focusing only on the inattentiveness of producers is that we get a two equation model for inflation and the output gap, which is sufficiently tractable to allow us to endogenise inattention and to introduce adaptive learning. We will start with a presentation of the model. Then we discuss the solution of the model under rational expectations and the conditions for uniqueness of the rational expectations equilibrium. Finally we endogenise the inattention parameter following Branch et al. (2004). The subsequent section will then investigate whether the model with endogenous inattention and adaptively learning agents is able to replicate the decline in the volatilities of inflation and the output gap after a break in monetary policy comparable to the one experienced in the US during the 1980s.

Note that nominal variables will be denoted by upper case letters and real variables by lower case letters. Denoting the aggregate price index by P_t, a nominal variable V_t and the corresponding real variable v_t are related via $v_t = V_t/P_t$. Logarithmic deviations from steady states will be marked with a hat.

3.3.1. Households. The typical household maximises his expected discounted lifetime utility

$$E_0\Big[\sum_{t=0}^{\infty}\delta^t u(c_t,l_t)\Big], \quad \delta\in(0,1), \tag{134}$$

where E_0 is the conditional expectation as of time 0, δ is the rate of time preference and $u(\cdot)$ is the per period utility given by

$$u(c_t,l_t) = \frac{c_t^{1-\sigma_c}-1}{1-\sigma_c} - \gamma\frac{l_t^{1+\sigma_l}+1}{1+\sigma_l}, \quad \gamma\geq 0. \tag{135}$$

The household derives utility from consumption c_t and leisure $1-l_t$, where $l_t\in(0,1)$ denotes hours worked. The utility function is assumed to be additively separable in both arguments and of the constant elasticity type. The parameters σ_c and σ_l are the inverse intertemporal elasticities of substitution of consumption and leisure. The maximisation is subject to the budget constraint

$$B_t + P_t c_t + P_t\tau_t \leq P_t w_t l_t + R_{t-1}B_{t-1} + \Omega_t. \tag{136}$$

The household receives income from labour $P_t w_t l_t$, dividends Ω_t from the firms and gross interest payments $R_t B_{t-1}$ from nominal bond holdings. It pays taxes for $P_t \tau_t$ and consumption $P_t c_t$ and invests in end of period nominal bond holdings B_t. Denoting inflation by $\pi_t = P_t / P_{t-1}$ the budget constraint in real terms is

$$b_t + c_t + \tau_t = w_t l_t + \pi_t^{-1} R_{t-1} b_{t-1} + \Omega_t / P_t. \tag{137}$$

The FOCs and the transversality condition of the household's problem are

$$w_t c_t^{-\sigma_c} = \gamma l_t^{\sigma_l}, \tag{138}$$
$$c_t^{-\sigma_c} = \delta R_t E_t[c_{t+1}^{-\sigma_c} / \pi_{t+1}], \tag{139}$$
$$0 = \lim_{t \to \infty} \delta^t u_{c_t} b_t. \tag{140}$$

Aggregate consumption c_t is a Dixit-Stiglitz aggregator of the consumption of a variety of goods $i \in (0,1)$

$$c_t \equiv \Big(\int_0^1 c_{it}^{(\varepsilon_t - 1)/\varepsilon_t} di \Big)^{\varepsilon_t / (\varepsilon_t - 1)}.$$

The parameter ε_t is the elasticity of substitution between any two goods at time t. It is assumed to follow an exogenous stochastic process. Cost minimisation on part of the households implies the optimal demand $c_{it} = (P_{it}/P_t)^{-\varepsilon_t} c_t$ for good i, where P_{it} is the price of good i and the aggregate price index P_t is

$$P_t = \Big(\int_0^1 P_{it}^{(1-\varepsilon_t)} di \Big)^{1/(1-\varepsilon_t)}. \tag{141}$$

3.3.2. Government. Monetary policy sets the nominal rate of interest R_t according to a Taylor rule

$$R_t = \phi_\pi \log \pi_t + \phi_z (\log y_t - \log y_t^p). \tag{142}$$

where $\pi_t = \Delta P_t$ is the inflation rate and y_t^p refers to potential output. [2] Potential output y_t^p is defined as the efficient level of output, which is the equilibrium level of output under complete information and perfect competition. The government also purchases an aggregate of individual goods g_t, which is of the same form as the consumption aggregator of households

$$g_t \equiv \Big(\int_0^1 g_{it}^{(\varepsilon_t - 1)/\varepsilon_t} di \Big)^{\varepsilon_t / (\varepsilon_t - 1)}.$$

[2] Arguably there may be persistent deviations from the rule (142), cf. e.g. Rudebush (2002). However, the introduction of a monetary policy shock would merely add another source of variation to the IS curve. In this sense it is equivalent to a government expenditure shock.

Optimisation by the government leads to the demand $g_{it} = (P_{it}/P_t)^{-\varepsilon_t} g_t$ for good i. Market clearing requires that for each good i, $y_{it} = c_{it} + g_{it}$. Defining aggregate output as

$$y_t \equiv \Big(\int_0^1 y_{it}^{(\varepsilon_t - 1)/\varepsilon_t} di \Big)^{\varepsilon_t/(\varepsilon_t - 1)},$$

it follows that the aggregate resource constraint can be written as

$$y_t = c_t + g_t. \tag{143}$$

Government expenditures g_t are assumed to follow an exogenous stochastic process. To finance government expenditures $P_t g_t$ the government issues nominal bonds B_t and collects taxes $P_t \tau_t$ from the household. It also repays bond holdings $R_t B_{t-1}$ of the previous period including interest. Thus the flow budget constraint of the government is

$$B_t + P_t \tau_t = P_t g_t + R_{t-1} B_{t-1}. \tag{144}$$

Government bonds B_{t+i} are subject to the solvency constraint $\lim_{i \to \infty} (B_{t+i} \prod_{s=0}^{i-1} R_{t+s}^{-1}) = 0$. In real terms the budget constraint can be rewritten as

$$b_t + \tau_t = \pi_t^{-1} R_{t-1} b_{t-1} + g_t. \tag{145}$$

3.3.3. Firms. There is a continuum $i \in (0,1)$ of monopolistically competitive firms corresponding to each good i. Each firm produces y_{it} units of the differentiated good i using specialised labour l_{it}. All firms share the same technology

$$y_{it} = a_t l_{it}^{\beta}, \tag{146}$$

where a_t is aggregate productivity and β measures the degree of returns to scale. We assume that there is a fixed stock of capital, so that capital can be suppressed in the production function and we can interpret β as the labour share. Aggregate productivity a_t follows an exogenous stochastic process. Hours worked are aggregated according to $l_t = \int_0^1 l_{it} di$.

Following Mankiw and Reis (2006a) and Mankiw and Reis (2006b) we assume that information is sticky. Each period a randomly drawn fraction λ of firms update their information. A firm that has last updated its information j periods ago will choose its nominal price P_{jt} to maximise expected real profits

$$\max_{P_{jt}} E_{t-j} \big[P_{jt}/P_t y_{jt} - w_t l_{jt} \big]$$

subject to the demand for good j

$$y_{jt} = (P_{jt}/P_t)^{-\varepsilon_t} y_t \tag{147}$$

and the production technology (146). [3] The first order condition is

$$P_{jt} = \frac{E_{t-j}[\varepsilon_t w_t l_{jt}]}{E_{t-j}[\beta(\varepsilon_t - 1) y_{jt}/P_t]}. \tag{148}$$

The main focus throughout the chapter will lie on the determination of the updating probability λ. A high λ means firms update their information sets relatively often, whereas a low level of λ corresponds to relatively infrequent updates of the information sets. A firm which has not updated for a while cannot react to all shocks that have realised since its last update. Therefore, all other things being equal, we would expect a decrease in the updating probability λ to lead to a reduction in the volatility of the price level. The updating probability λ is often referred to as the 'inattention parameter' or 'information acquisition rate'.

3.3.4. The Log-Linearised Model. We start with the computation of the efficient level of output, which is the equilibrium level of output under complete information and perfect competition in the absence of nominal rigidities. Under complete information (148) can be rewritten as

$$\frac{P_{jt}}{P_t} = \mu_t \frac{w_t l_{jt}}{\beta y_{jt}}. \tag{149}$$

where μ_t is the markup given by $\mu_t = \varepsilon_t/(\varepsilon_t - 1)$. The system of equations (146), (147) and (149) yields a unique solution for the individual output y_{jt}. Thus under complete information all firms will choose the same level of output. The demand function (147) then implies a relative price P_{jt}/P_t equal to unity. It follows that the efficient level of output y_t^p is implicitly defined by the relation

$$1 = \frac{w_t (y_t^p/a_t)^{1/\beta}}{\beta y_t^p}. \tag{150}$$

We are now ready to compute a log-linear approximation of the model around its non-stochastic steady state. Log-linearisation of the household's first order conditions (138) and (139) yields

$$\hat{w}_t = \sigma_c \hat{c}_t + \sigma_l \hat{l}_t, \tag{151}$$

$$\hat{c}_t = E_t \hat{c}_{t+1} - \sigma_c^{-1}(\hat{R}_t - E_t \hat{\pi}_{t+1}). \tag{152}$$

[3] There is a slight abuse of notation because there is actually a continuum of firms with mass $\lambda(1-\lambda)^j$, who all set a price P_{it} for good i based on information, which is j periods old. However, they all solve the same problem based on the same information, so we do not need to distinguish them.

The resource constraint (143) and the aggregate of the production function (146) can be approximated as [4]

$$\begin{aligned} \hat{y}_t &= c/y\hat{c}_t + \hat{g}_t, && (153) \\ \hat{y}_t &= \hat{a}_t + \beta\hat{l}_t. && (154) \end{aligned}$$

Using the resource constraint (153) and the production function (154) we can rewrite the expression for the real wage (151) as

$$\hat{w}_t = (\sigma_y + \sigma_l/\beta)\hat{y}_t - \sigma_y\hat{g}_t - \sigma_l/\beta\hat{a}_t, \quad (155)$$

$\sigma_y \equiv y/c\sigma_c$ is the inverse intertemporal elasticity of substitution of aggregate expenditure. Provided the distortions from the markup are sufficiently small we can approximate the efficient output (150) by [5]

$$\hat{y}_t^p = \Xi_a\hat{a}_t + \Xi_g\hat{g}_t, \quad (156)$$

with

$$\Xi_a = \frac{1+\sigma_l}{\beta(\sigma_y - 1) + \sigma_l + 1} \quad \text{and} \quad \Xi_g = \frac{\beta\sigma_y}{\beta(\sigma_y - 1) + \sigma_l + 1}. \quad (157)$$

The real wage has been substituted out using (155). Using the production function (146) and the demand function (147) to substitute out hours worked l_{jt} and the demand y_{jt} for good j, the log-linearised first order condition (148) of the firms' pricing problem is

$$\hat{P}_{jt} = E_{t-j}\left[\hat{P}_t + \frac{\beta\hat{w}_t + (1-\beta)\hat{y}_t - \hat{a}_t - \beta/(\varepsilon-1)\hat{\varepsilon}_t}{\beta + \varepsilon(1-\beta)}\right]. \quad (158)$$

Up to a first order approximation the price index (141) equals $\hat{P}_t = \lambda\sum_{j=0}^{\infty}(1-\lambda)^j\hat{P}_{jt}$. Let $z_t = y_t - y_t^p$ be the output gap. Combining the price index with the expressions for the optimal price (158), the real wage (155) and potential output (150) we have

$$\hat{P}_t = \lambda\sum_{j=0}^{\infty}(1-\lambda)^j E_{t-j}\left[\hat{P}_t + \alpha\hat{z}_t + \hat{u}_t\right], \quad (159)$$

with $\alpha = (\beta(\sigma_y - 1) + \sigma_l + 1)/(\beta + \varepsilon(1-\beta))$ and $\hat{u}_t = -\beta/[(\varepsilon-1)(\beta+\varepsilon(1-\beta))]\hat{\varepsilon}_t$. Note that due to our assumption of sticky information the current price level depends on all past expectations of the current price level, the output gap and the shock $\hat{u}_t$. The shock $\hat{u}_t$ is often called a cost push shock. Although we have derived it as a shock on the markup of the producers, market power in the labour market or time-varying taxes would enter the model in the same way.

[4] For notational convenience we follow the standard practice and let $\hat{g}_t$ denote the deviation of g_t from its steady state value expressed as a percentage of the steady state value of aggregate output y_t, i.e. $\hat{g}_t = (g_t - g)/y$.

[5] For details of the approximation confer the discussion in Woodford (2003b) about the 'small Φ_y approximation'.

For $0 < \lambda < 1$ the equation for the price level can be transformed in a relationship for the inflation rate $\hat{\pi}_t = \Delta\hat{P}_t$, which has been labelled the sticky information Philips curve (cf. also Mankiw and Reis, 2002) [6]

$$\hat{\pi}_t = \frac{\lambda}{1-\lambda}(\alpha\hat{z}_t + \hat{u}_t) + \lambda \sum_{j=0}^{\infty}(1-\lambda)^j E_{t-1-j}[\hat{\pi}_t + \alpha\Delta\hat{z}_t + \Delta\hat{u}_t]. \tag{160}$$

Combining the intertemporal Euler equation (152) with the aggregate resource constraint (153) and the equation for potential output (150) we obtain the intertemporal IS curve

$$\hat{z}_t = E_t\hat{z}_{t+1} - \sigma_y^{-1}(\hat{R}_t - E_t\hat{\pi}_{t+1}) + \hat{v}_t. \tag{161}$$

where $\hat{v}_t = E_t[\Delta\hat{y}^p_{t+1} - \Delta\hat{g}_{t+1}] = E_t[\Xi_a\Delta\hat{a}_{t+1} - (1-\Xi_g)\Delta\hat{g}_{t+1}]$. Due to consumption smoothing the output gap today depends positively on the output gap tomorrow and negatively on the real rate of interest $\hat{R}_t - E_t\hat{\pi}_{t+1}$. The demand shock $\hat{v}_t$ is a composite shock which reflects predictable changes in potential output and government expenditures. A predictable rise in productivity raises the current output gap just as an anticipated fall in government expenditures ($\Xi_g < 1$). Other possible sources for the demand shock that would enter the model in the same way as the government expenditure shock are fluctuations in the household's preferences or a monetary policy shock.

The final model consists of the sticky information Philips curve (160), the intertemporal IS curve (161) and the log-linearisation of the Taylor rule

$$\hat{R}_t = \phi_\pi\hat{\pi}_t + \phi_y\hat{z}_t. \tag{162}$$

For the exogenous processes of the cost push shock and the demand shock we assume that they follow the AR(1) processes

$$\hat{u}_t = \rho_v\hat{u}_{t-1} + e_{u,t} \quad \text{and} \quad \hat{v}_t = \rho_v\hat{v}_{t-1} + e_{v,t}, \tag{163}$$

where $e_{u,t}$ and $e_{u,t}$ are white noise with standard deviations σ_u and σ_v respectively.

Under rational expectations the nominal rate of interest can be substituted out of the IS curve using the Taylor rule. The resulting reduced form IS curve will be needed to compute the rational expectations equilibrium. Combining (161) and (162) we get the reduced IS curve

$$\hat{z}_t = (\sigma_y + \phi_y)^{-1}(\sigma_y E_t\hat{z}_{t+1} - \phi_\pi\hat{\pi}_t + E_t\hat{\pi}_{t+1} + \sigma_y\hat{v}_t). \tag{164}$$

[6] Under full attention ($\lambda = 1$) equation (159) simplifies to $z_t = -u_t/\alpha$.

3.4. Equilibrium Concepts

3.4.1. Rational Expectations Equilibrium. A remarkable insight of Wang and Wen (2006) is that the introduction of sticky information does not change the determinacy properties of a log-linear DSGE model. Thus it suffices to check determinacy under the assumption that all agents have complete information. It is then easy to check that the model is determinate if and only if the Taylor principle holds (ϕ_π>1). Thus, in contrast to a New Keynesian model with sticky prices, output gap stabilisation (i.e. $\phi_z > 0$) does not contribute to equilibrium uniqueness. [7] The intuition for this result is as follows. In the New Keynesian model with sticky prices, a permanent increase in inflation is accompanied by a permanent increase in the output gap. Therefore, a tightening of monetary policy following an increase in the output gap reduces the danger of self-fulfilling inflation. In the sticky information version of the model this not true. Rearranging equation (160) we have

$$\hat{z}_t = (1-\lambda)\hat{z}_{t-1} + \frac{\lambda}{\alpha}\sum_{j=1}^{\infty}(1-\lambda)^{j+1}[\hat{\pi}_t + \Delta\hat{z}_t - E_{t-j}(\hat{\pi}_t + \Delta\hat{z}_t)] + \hat{w}_t, \qquad (165)$$

where $\hat{w}_t$ is some exogenous stochastic process. Obviously, a permanent increase in inflation will increase both inflation and expected inflation and leave the output gap unchanged.

To compute the REE we can write the aggregate price index and the output gap as a function of all past disturbances and use the method of undetermined coefficients. Define $\Lambda_n = \lambda\sum_{i=0}^{n}(1-\lambda)^i$, which is the size of agents that have updated their information sets within the last n periods. Moreover let $s \in S = \{u, v\}$ denote the cost push shock and the demand shock and let $\mathbb{1}_{s=v}(s)$ be the indicator function of the demand shock, which equals unity if $s = v$ and zero otherwise. For $0 < \lambda \leq 1$ the rational expectations equilibrium can now be computed using the following proposition.

PROPOSITION 3.4.1. *Assume that* $0 < \lambda \leq 1$. *Let* $s \in S = \{u, v\}$ *denote the cost push shock and the demand shock. Let* $p_n(s)$ *or* $z_n(s)$ *be the impact of shock* s *at lag* n *on prices or the output gap respectively. Then*

$$\hat{P}_t = \sum_{s\in S}\sum_{n=0}^{\infty} p_n(s)e_{s,t-n}, \quad \hat{z}_t = \sum_{s\in S}\sum_{n=0}^{\infty} z_n(s)e_{s,t-n}. \qquad (166)$$

Define the parameters

$$A_n = 1 + \frac{\sigma_y}{\alpha}\frac{1-\Lambda_n}{\Lambda_n} \quad B_n = 1 + \phi_\pi + \frac{\sigma_y + \phi_y}{\alpha}\frac{1-\Lambda_n}{\Lambda_n} \qquad (167)$$

[7] Recall that the analogous condition with a sticky price version of the Philips curve $\pi_t = \kappa z_t + \beta E_t \pi_{t+1}$ is that $\phi_\pi + (1-\beta)/\kappa\phi_z > 1$ (cf. e.g. Woodford, 2003b).

and

$$C_n(s) = \begin{cases} (\sigma_y(1-\rho_u)+\phi_y)/\alpha\rho_u^n & \text{for} \quad s=u \\ \sigma_y\rho_v^n & \text{for} \quad s=v \end{cases} . \tag{168}$$

For $s \in S = \{u,v\}$ the coefficients of the law of motion of the price level $p_n(s)$ solve the second order difference equation

$$A_{n+1}p_{n+1}(s) - B_n p_n(s) + \phi_\pi p_{n-1}(s) = -C_n(s). \tag{169}$$

with the boundary conditions $p_{-1} = 0$ and $\lim_{n\to\infty}(p_n - p_{n-1}) = 0$. The coefficients of the law of motion of the output gap are

$$z_n(s) = \alpha^{-1}\left[(1-\Lambda_n)/\Lambda_n p_n(s) - \mathbb{1}_{s=u}(s)\rho_u^n\right]. \tag{170}$$

Equation (170) is a direct consequence of the method of undetermined coefficients applied to the aggregate supply relation (159). Application of the same method to the IS curve (164) and using equation (170) to substitute out the coefficients of the output gap yields the difference equation (169). To solve (169) we use the same approach as Mankiw and Reis (2006b). First we choose some large enough N and set $C_N = 0$. Then we numerically solve the linear system of equations (169) for $n = 0,\ldots,N$ with $p_{-1} = 0$ and $p_{N-1} = p_N$. The output gap coefficients follow from (170).

3.4.2. Endogenous Inattention. In the previous paragraph we have derived the rational expectations equilibrium of the model taking the inattention parameter as given. In this paragraph we follow Branch et al. (2006) and let agents endogenously choose their information acquisition rate λ. Recall that in equilibrium both the price level $P_t(\bar{\lambda})$ and the output gap $z_t(\bar{\lambda})$ are functions of the economy wide information acquisition rate, which we subsequently denote by $\bar{\lambda}$. From equation (158) we see that the optimal price of a producer who has updated j periods ago is

$$\hat{P}_{jt}(\bar{\lambda}) = E_{t-j}[\hat{P}_t(\bar{\lambda}) + \alpha\hat{z}_t(\bar{\lambda}) + \hat{u}_t]. \tag{171}$$

Let $P_t^*(\lambda,\bar{\lambda})$ be the price charged by a specific firm, which updates its information each period with probability λ given that the economy wide updating probability equals $\bar{\lambda}$. Let s_t be a signal which takes on the value $s_t = 1$ if the firm updates its information in period t and $s_t = 0$ if not. Then the prices charged by the firm will evolve according to

$$P_t^*(\lambda,\bar{\lambda}) = \begin{cases} \hat{P}_{0t}(\bar{\lambda}) & \text{if} \quad s_t = 1 \\ \hat{P}_{jt}(\bar{\lambda}) & \text{if} \quad s_{t-j} = 1,\ s_{t-j+1} = \ldots = s_t = 0 \end{cases} , \tag{172}$$

where $\lambda = \text{Prob}(s_t = 1)$. [8] Following Branch et al. (2006) we assume that producers choose their information acquisition rate λ to minimise the unconditional mean squared error between their price $P_t^*(\lambda, \bar{\lambda})$ and the price under complete information $P_t^*(1, \bar{\lambda}) = \hat{P}_{0t}(\bar{\lambda})$ given that information is costly. This leads to the definition of the loss function

$$L(\lambda, \bar{\lambda}) = MSE(\lambda, \bar{\lambda}) + C\lambda^2, \tag{173}$$

where $MSE(\lambda, \bar{\lambda}) = E(\hat{P}_{0t}(\bar{\lambda}) - P_t^*(\lambda, \bar{\lambda}))^2$ is the unconditional mean squared error. We assume that the process of acquiring and processing information is costly ($C > 0$). The best response function of a specific firm maps all aggregate levels of inattention $\bar{\lambda}$ into an individually optimal value of inattention

$$B(\bar{\lambda}) = \arg \min_{\lambda \in (0,1]} L(\lambda, \bar{\lambda}). \tag{174}$$

The following proposition establishes that the mean squared error made by the producers is strictly decreasing and convex in λ. It follows that the minimum in (174) exists because the mean squared error is strictly decreasing in λ, whereas the marginal costs of updating $2C\lambda$ vanish as λ goes to zero. Therefore it is optimal to choose $\lambda > 0$. The fact that the mean squared error is a strictly convex function of λ implies that the loss function $L(\lambda, \bar{\lambda})$ is also strictly convex in λ so that the minimum in (174) is unique for all $\bar{\lambda}$.

PROPOSITION 3.4.2. *The mean squared error $MSE(\lambda, \bar{\lambda})$ is a strictly decreasing and convex function of λ.*

PROOF. Obviously it is sufficient to show that the mean squared error is a convex function of λ. From the definition of the optimal price under sticky information (172) it follows that

$$\hat{P}_{0t}(\bar{\lambda}) - P_t^*(\lambda, \bar{\lambda}) = \begin{cases} 0 & \text{if} \quad s_t = 1 \\ \hat{P}_{0t}(\bar{\lambda}) - \hat{P}_{jt}(\bar{\lambda}) & \text{if} \quad s_{t-j} = 1,\ s_{t-j+1} = \ldots = s_t = 0 \end{cases}. \tag{175}$$

Combining equations (166) and (171) yields the MA(∞) representation

$$\hat{P}_{jt}(\bar{\lambda}) = \sum_{s \in S} \sum_{n=j}^{\infty} (\bar{p}_n(s) + \alpha \bar{z}_n(s) - c_n(s)) e_{s,t-n} \tag{176}$$

where

$$c_n(s) = \begin{cases} \dfrac{\beta}{(\varepsilon - 1)(\beta + \varepsilon(1-\beta))} \rho_u^n & \text{if} \quad s = \varepsilon \\ 0 & \text{if} \quad s = v \end{cases}. \tag{177}$$

[8] Note that a deviation of a single firm off equilibrium has no effect upon the aggregate price level or the output gap because there is a continuum of firms. The optimal price of a firm in turn depends on the level of inattention only indirectly through its effect on the price level and the output gap. Therefore the optimal price of a firm which plans to deviate from the aggregate level of inattention does not depend the inattentiveness that it plans to choose, i.e. a firm with information as of $t - j$ will set a price equal to $\hat{P}_{jt}(\bar{\lambda})$.

The bar over $\bar{p}_n$ and $\bar{z}_n$ indicates the implicit dependence on $\bar{\lambda}$. It follows that

$$\hat{P}_{0t}(\bar{\lambda}) - \hat{P}_{jt}(\bar{\lambda}) = \sum_{s\in S}\sum_{n=0}^{j-1}(\bar{p}_n(s) + \alpha\bar{z}_n(s) - c_n(s))e_{s,t-n}. \tag{178}$$

Combining the law of motion (175) and the representation (178) of the forecast error we get the expression for the mean squared error

$$\begin{aligned} E(\hat{P}_{0t}(\bar{\lambda}) - P_t^*(\lambda,\bar{\lambda}))^2 &= \sum_{s\in S}\lambda\sum_{j=1}^{\infty}(1-\lambda)^j\sum_{n=0}^{j-1}(\bar{p}_n(s) + \alpha\bar{z}_n(s) - c_n(s))^2\sigma_s^2 \\ &= \sum_{s\in S}\sum_{n=0}^{\infty}(1-\lambda)^{n+1}(\bar{p}_n(s) + \alpha\bar{z}_n(s) - c_n(s))^2\sigma_s^2. \end{aligned} \tag{179}$$

From (179) it follows immediately that $\partial MSE/\partial\lambda < 0$ and $\partial^2 MSE/\partial\lambda^2 > 0$. □

Taking the aggregate information acquisition rate and the resulting processes for the price level and the output gap as given an individual firm minimises its loss over its individual information acquisition rate. The fixed point of this process is a Nash equilibrium, which we label endogenous inattention. [9]

DEFINITION 3.4.3. Endogenous inattention is a symmetric Nash equilibrium defined by the fixed point $\lambda^* = B(\lambda^*)$.

The following proposition establishes that endogenous inattention always exists. However the Nash equilibrium is not necessarily unique even though this will be the case in the calibration below.

PROPOSITION 3.4.4. *Endogenous inattention exists.*

PROOF. The proof is complicated slightly by the fact that the price index is not well defined for $\bar{\lambda} = 0$. Consider the limit

$$\begin{aligned} \lim_{\bar{\lambda}\to 0} B(\bar{\lambda}) &= \lim_{\bar{\lambda}\to 0} \arg\min_{\lambda\in(0,1]} L(\lambda,\bar{\lambda}) \\ &= \arg\min_{\lambda\in(0,1]}\sum_{s\in S}\sum_{n=0}^{\infty}(1-\lambda)^{n+1}(\lim_{\bar{\lambda}\to 0}\bar{p}_n(s) + \alpha\lim_{\bar{\lambda}\to 0}\bar{z}_n(s) - c_n(s))^2\sigma_s^2 + C\lambda^2 \end{aligned}$$

With $\lim_{\bar{\lambda}\to 0}\bar{p}(u) = \lim_{\bar{\lambda}\to 0}\bar{z}(u) = \lim_{\bar{\lambda}\to 0}\bar{p}(v) = 0$ and $\lim_{\bar{\lambda}\to 0}\bar{z}(v) = \frac{\sigma_y}{\sigma_y(1-\rho_y)+\phi_y}\rho_v^n$, it follows that

$$\lim_{\bar{\lambda}\to 0} B(\bar{\lambda}) = \arg\min_{\lambda\in(0,1]}\frac{(1-\lambda)k_1}{1-\rho_v^2(1-\lambda)} + \frac{(1-\lambda)k_2}{1-\rho_u^2(1-\lambda)} + C\lambda^2 > 0,$$

[9] The definition is taken from Branch et al. (2004).

where k_1 and k_2 are positive constants. The minimum exists because

$$\lim_{\lambda \to 0} \frac{\partial}{\partial \lambda} \left(\frac{(1-\lambda)k_1}{1-\rho_v^2(1-\lambda)} + \frac{(1-\lambda)k_2}{1-\rho_u^2(1-\lambda)} + C\lambda^2 \right) < 0.$$

Consider the function $f : (0,1] \to (0,1]$ defined by $f(\bar{\lambda}) = B(\bar{\lambda}) - \bar{\lambda}$. We have just established that $f(\bar{\lambda}) > 0$ for some $\bar{\lambda} > 0$. We assume that $f(1) < 0$, since otherwise there is nothing to prove. Noting that $f(\bar{\lambda})$ is continuous it follows from the intermediate value theorem that B has at least one fixed point. □

3.5. Adaptive Learning

The main focus of this chapter is on the interaction of adaptive learning and endogenous inattention. Under adaptive learning the timing is as follows. First the agents in the economy who have updated their information sets use the data on prices and the output gap that has realised up to the current period to estimate a new forecasting model. Then they use the estimated processes for the data to choose the currently optimal inattention parameter. Then all agents use their current forecasting models to form their expectations. Finally, the expectations and the chosen inattention parameters determine the new realisations of prices and the output gap.

The first step in the learning process is the estimation of a perceived relationship for inflation and the output gap. We assume that the perceived laws of motion of inflation and the output gap are of the same functional form as the rational expectations equilibrium representation (166). Thus, each period, the agents who update their information sets estimate two moving average processes of the form

$$\hat{P}_t = \sum_{s \in S} \sum_{n=0}^{\infty} \hat{p}_n(s) e_{s,t-n}, \quad \hat{z}_t = \sum_{s \in S} \sum_{n=0}^{\infty} \hat{z}_n(s) e_{s,t-n}, \tag{180}$$

where $\hat{p}_n(s)$ and $\hat{z}_n(s)$ is the estimated impact of shock s at lag n on prices and the output gap. Recall that $s \in \{u, v\}$ is a placeholder for the markup and the demand shock. [10]

Each period the fraction of producers who update their information set also solve for the optimal information acquisition rate λ_t. In the learning literature it is standard to separate forecasting and decision problems. In this tradition we assume that the firms take the estimated perceived law of motion as given when choosing the optimal inattentiveness. Using the expression (179) for the mean squared error the estimated loss function is

$$L(\lambda) = \sum_{s \in S} \sum_{n=0}^{\infty} (1-\lambda)^{n+1} (\hat{p}_n(s) + \alpha \hat{z}_n(s) - c_n(s))^2 \sigma_s^2 + C\lambda^2. \tag{181}$$

[10] We follow the standard practice in the learning literature and assume that the agents can observe the shocks. Therefore the coefficients of the perceived law of motion can be estimated recursively using ordinary least squares.

where $c_n(s)$ is the exponentially decaying series given by (177). Taking the estimates of the perceived law of motion as given the currently optimal rate of inattention can be computed as

$$\lambda_t = \arg\min_{\lambda \in (0,1]} L(\lambda). \tag{182}$$

Proposition 3.4.2 applies analogously, so that λ_t is uniquely defined. Based on their most recent estimates, the agents compute forecasts of the price level, the output gap, the inflation rate and output gap growth, i.e. they calculate

$$\forall j > 0: \quad E_t P_{t+j}, \quad E_t z_{t+j}, \quad E_t \pi_{t+j}. \tag{183}$$

Finally, given the expectations of the agents the actual law of motion of the price level and the output gap are

$$\hat{P}_t = \sum_{j=0}^{\infty} w_{t,j} E_{t-j}[\hat{P}_t + \alpha \hat{z}_t + \hat{u}_t] \tag{184}$$

$$\hat{z}_t = E_t \hat{z}_{t+1} - \sigma_y^{-1}(\hat{R}_t - E_t \hat{\pi}_{t+1}) + \hat{v}_t \tag{185}$$

where the weights $w_{t,j}$ evolve according to

$$w_{t,j} = \begin{cases} (1-\lambda_{t-j}) w_{t-1,j-1} & \text{if} \quad j > 1 \\ \sum_{i=0}^{\infty} \lambda_{t-i} w_{t-1,i} & \text{if} \quad j = 0 \end{cases}. \tag{186}$$

The weight $w_{t,j}$ reflects the proportion of producers at time t who last updated their information set j periods ago. Under learning the weights become history dependant. The fraction of producers who last updated their information sets in period $t-j$ will stick to the optimal information acquisition rate λ_{t-j} at the time of their last update. So the size of agents with information from period $t-j$ will decay at the rate $1-\lambda_{t-j}$. Provided λ_t converges to a fixed value λ_∞, the weights converge to $w_{\infty,j} = \lambda_\infty(1-\lambda_\infty)^j$ so that eventually the price level under learning (159) converges to the price level under rational expectations (184) with $\lambda = \lambda_\infty$. [11]

The following section presents a simulation of a calibrated instance of the model. As in the computation of the rational expectations equilibrium of the model we have to truncate all infinite sums at some number N. For sufficiently large N the particular choice of the truncation point seems to be irrelevant.

[11] Note that Branch et al. (2006) use a different definition for the weights, which is incorrect in the transition to the equilibrium. Especially it can be checked that the weights in their definition do not necessarily sum up to one.

3.6. Simulation Results

We will now conduct a simulation study, which is intended to check in how far the small learning model we have constructed is capable to produce a simultaneous decline of the volatility in inflation and the output gap. We set the intertemporal elasticity of substitution of output to $\sigma_y = 1$ and the sensitivity of the firm's relative price with respect to changes in the output gap to $\alpha = 0.1$. The parameters of the Taylor rule are set to $\phi_{\pi,1} = 1.1$ and $\phi_{y,1} = 0.1$ before the break and $\phi_{\pi,2} = 1.5$ and $\phi_{y,2} = 0.2$ thereafter. For the stochastic processes we set the AR(1) parameters to $\rho_u = 0.9$ for the cost push shock and $\rho_v = 0.95$ for the demand shock. The standard deviations of the shocks are calibrated to $\sigma_u = 0.23$ and $\sigma_v = 0.05$. Finally we choose information costs to equal $C = 25$. We deliberately choose a coefficient $\phi_{\pi,1} > 1$ for the period before the break in order to obtain a determinate equilibrium both before and after the break in monetary policy. Many authors have estimated a feedback coefficient ϕ_π for inflation in the Taylor rule, which is slightly below one (cf. Clarida et al., 2000). Yet, we choose a value which is marginally above one because we want to abstract from any changes in the volatilities due to sunspots.

The calibration implies a Nash equilibrium of $\bar{\lambda}_1 = 0.41$ for the period before the break and a value of $\bar{\lambda}_2 = 0.27$ thereafter. Thus on average the producers update their information set every two and a half periods before and every four periods after the break. The value after the break is in line with the estimates of Mankiw and Reis (2002) and Khan and Zhu (2002). Figure 3 plots the best response function of the producers under both regimes. There is a unique Nash equilibrium under both regimes. Moreover both equilibria are globally stable in the sense that $B' < 1$.

The calibration has been chosen to replicate closely the volatility of inflation and the output gap we have found in the data. Table 1 compares the standard deviations of the data to the equilibrium values for the calibrated model. The column 'fixed inattention' reports the volatilities for the case where the inattentiveness of the producers is exogenously fixed at the level before the break. In the data the standard deviation of inflation drops by 72% – the standard deviation of the output gap by 49%. Keeping the inattentiveness fixed the standard deviation in the model fall by 46% in case of inflation and rises by 18% in case of the output gap, reflecting the orthodox volatility trade-off. Under endogenous inattention the decline in the attentiveness of the producers leads to an overall decline in the standard deviations by 70% and 30% respectively. So, under rational expectations we would expect an increase in the output gap by 18% if we fix the inattentiveness exogenously. However the endogenous reaction in the information acquisition of the producers leads to an overall decline by 30% as suggested by the actual historical experience.

We simulate the model over $10,000$. We initialise the beliefs of the agents, i.e. the MA parameters of the reduced form they estimate and the initial inattention parameter at their

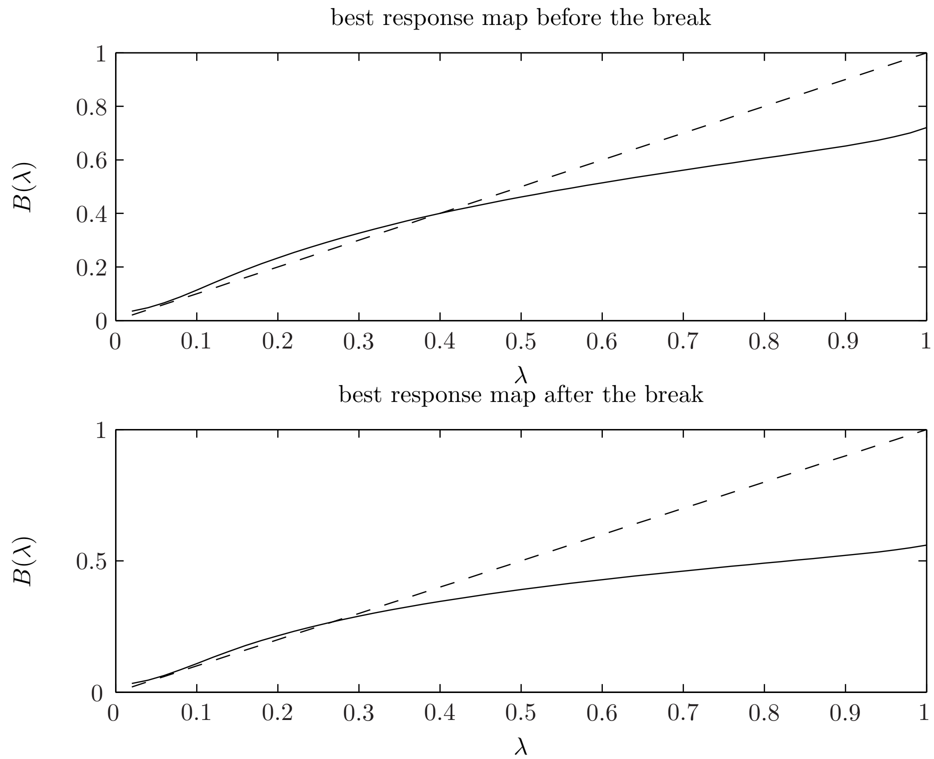

FIGURE 3. *best response functions before and after the break in monetary policy*

	data		model endo. inattention		model fixed inattention	
	σ_π	σ_z	σ_π	σ_z	σ_π	σ_z
before the break	2.67	2.99	2.29	2.28	2.29	2.28
after the break	0.76	1.52	0.67	1.6	1.23	2.77
change	-72%	-49%	-70%	-30%	-46%	+18%

TABLE 1. *standard deviations of the model and the actual data*

respective values under rational expectations. Agents learn using recursive least squares. In order to allow the agents to track structural changes in the model, we assume that they discount past data with a discount factor of 0.99. In period $2,000$ we change the parameters of the Taylor rule from $(\phi_{\pi,1}, \phi_{y,1})$ to $(\phi_{\pi,2}, \phi_{y,2})$ to simulate the break in monetary policy. Figure 4 plots the evolution of the inattention parameter under learning. The model seems to be stable under learning as the beliefs of the agents including the inattention parameter stay constant before the break. After the break the inattention parameter falls continuously until the new equilibrium value is reached to which it appears to converge.

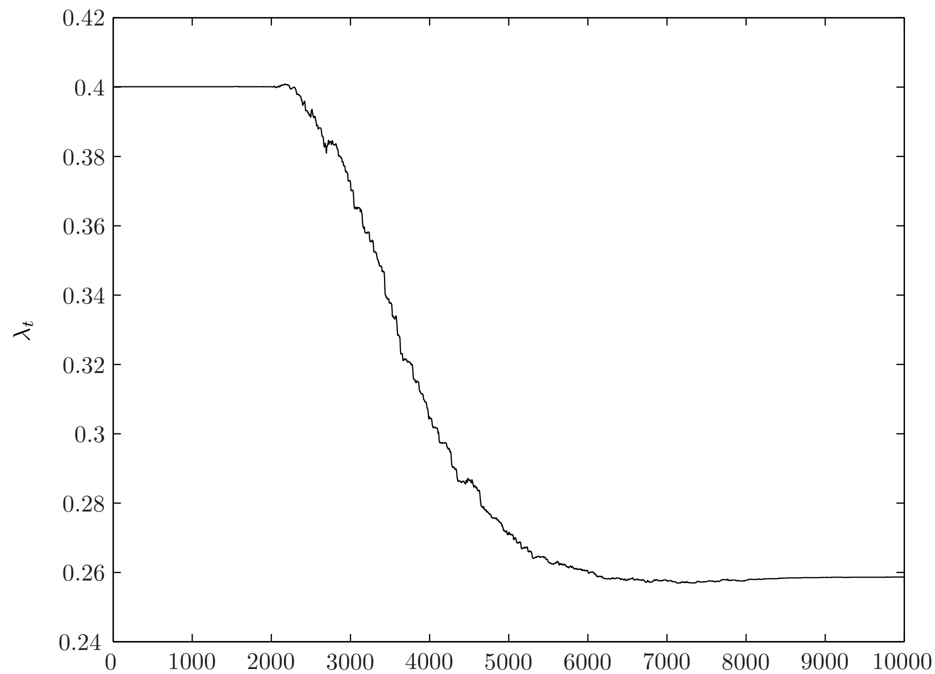

FIGURE 4. *evolution of the updating probability λ_t under learning*

Figure 5 plots the volatility of inflation and the output gap under learning (solid line) against the equilibrium volatilities with endogenous inattention (dashed line) and with exogenously fixed inattention (dotted line). In the case of exogenous inattention we compute the equilibrium values for the volatility of inflation and the output gap keeping the inattention parameter fixed at the value before the break. This benchmark shows which part of the volatilities is due to the break in monetary policy and which part is due to the endogenous change of the level of inattention of the producers in response to the policy break. In the case where the updating probability of the producers is exogenously given we see that there is a trade-off between the volatility of inflation and the output gap. The increased focus on inflation stabilisation will in equilibrium reduce the volatility of inflation but at the same time increase the volatility of the output gap. However the Nash equilibrium value for the updating probability decreases from $\bar{\lambda}_1 = 0.41$ before the break to $\bar{\lambda}_2 = 0.27$ after the break. Instead of updating every two and a half periods agents will on average only want to update every four periods. This is a result of the increased stability of inflation, which results in a lower volatility of the optimal price charged by the producers. As a consequence information will disseminate more slowly in the economy. Thus if the producers choose their information acquisition rate endogenously both volatilities will decrease compared to the exogenous case. This effect can be so large that the trade-off between the volatility of inflation and aggregate economic activity breaks down, which is what we see in the calibration.

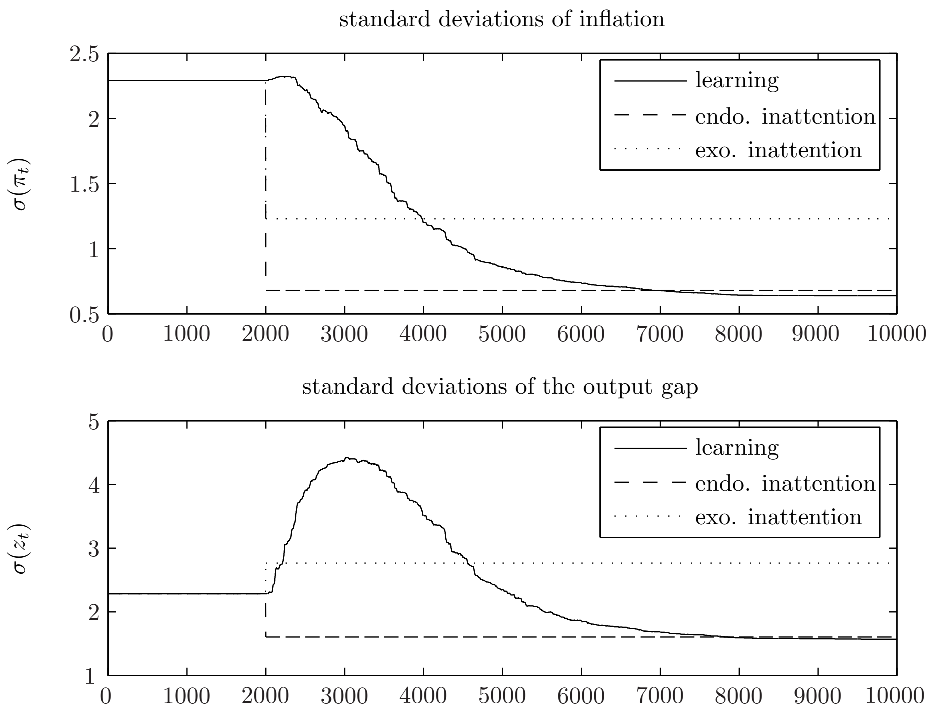

FIGURE 5. *standard deviations under learning compared to the equilibrium values in case of exogenous and endogenous inattention*

Under rational expectations we would see an immediate drop in both volatilities. The fact that we model learning as being adaptive introduces inertia in the beliefs of the agents. In the short run the inattention parameter does not change very much, so that the volatility of the output gap increases, whereas the volatility of the inflation rate falls rather quickly. The initial rise in the volatility of inflation leads to revisions in the beliefs of the agents that increase the volatility even further. Therefore the short run volatility trade-off is even steeper under learning than in the case of exogenous inattention. However, as new data arrives the agents revise their beliefs towards the new equilibrium so that in the long run the trade-off breaks down and both volatilities converge to their new equilibrium values. The magnitude of the rise of the volatility of the output gap in the short run and the speed of convergence to the new equilibrium depend on the details of the way agents learn, i.e. their perceived law of motion of the economy, the learning algorithm and the parameters of the learning algorithm. The qualitative result however is independent of the specific implementation of the learning algorithm. Adaptive learning introduces a short run trade-off between the stabilisation of inflation and the output gap, whereas in the long run the beliefs of the producers converge to the new equilibrium so that both volatilities decline.

3.7. Conclusions

We have set up a model of inflation and the output gap in which the information of producers about the economy is sticky. In order to set their prices optimally producers need to forecast aggregate economic variables like inflation and the output gap. We have assumed that the producers choose their information acquisition rate to balance the forecast error they make against the cost associated with the information acquisition. Monetary policy sets the nominal interest rate as a linear feedback rule of inflation and the output gap. In our calibration exercise we have shown that in the long run a shift of monetary policy towards inflation stabilisation may lead to a simultaneous decline in both the volatility of inflation and the output gap. The traditional trade-off between the stabilisation of inflation and the output gap eventually breaks down as the producers choose to update their information less frequently. The reason for them to update less frequently is that a reduction in inflation volatility also reduces the volatility of the price they would like to set if information were costless. Although in the long run monetary policy faces no trade-off, the inertia in the beliefs introduced by adaptive learning creates a volatility trade-off in the short run. The short run trade-off introduced by learning may even be steeper than the traditional trade-off in a model without learning and endogenous inattention.

Our results show that a model with sticky information, endogenous inattention and adaptively learning agents may produce a ‘great moderation’ as a consequence of a shift in monetary policy towards more inflation stabilisation. The coincidence of the break in US monetary policy under Fed chairman Paul Volcker and the break in the volatility of the output gap suggest that monetary policy may at least have contributed some part to the great moderation. A New Keynesian model with sticky information and adaptively learning agents gives an explanation how a shift in monetary policy can lead to a long run decline in both volatilities, while at the same time preserving a significant trade-off in the short run.

Appendix

A.1. Appendix to Chapter 1

A.1.1. First and Second Order Forms (Positive Bond Holdings). Under contemporaneous data rules the monetary policy feedback rule $R_t = \phi_\pi \hat{\pi}_t + \phi_z \hat{z}_t$ can be substituted in (21) to yield

$$\hat{b}_t \quad = \quad (\beta^{-1} - \gamma_1)\hat{b}_{t-1} - \beta^{-1}(\hat{\pi}_t - \phi_\pi \hat{\pi}_{t-1} - \phi_z \hat{z}_{t-1}) + \frac{g}{b}\hat{g}_t. \tag{A.1}$$

To obtain a second order form use (19), (20) and the interest rate rule $\hat{R}_t = \phi_\pi \hat{\pi}_t + \phi_z \hat{z}_t$ to solve for inflation

$$\hat{\pi}_t = \frac{\kappa + \beta(\sigma_y + \phi_z)}{\phi_z + \sigma_y + \phi_\pi \kappa} E_t \hat{\pi}_{t+1} + \frac{\kappa \sigma_y}{\phi_z + \sigma_y + \phi_\pi \kappa} E_t \hat{z}_{t+1} + \frac{\kappa \sigma_y \rho_g}{\phi_z + \sigma_y + \phi_\pi \kappa} \hat{g}_t, \tag{A.2}$$

where $\rho_g = g/y + \sigma_c/(\sigma_y + \sigma_l)g/c$. Substituting this result in $(A.1)$ yields

$$\begin{aligned} \hat{b}_t \quad &= \quad -\frac{\phi_z + \sigma_y + \beta^{-1}\kappa}{\phi_z + \sigma_y + \phi_\pi \kappa} E_t \hat{\pi}_{t+1} - \frac{\beta^{-1}\kappa \sigma_y}{\phi_z + \sigma_y + \phi_\pi \kappa} E_t \hat{z}_{t+1} + \beta^{-1}\phi_\pi \hat{\pi}_{t-1} \\ &+ \quad \beta^{-1}\phi_z \hat{z}_{t-1} + (\beta^{-1} - \gamma_1)\hat{b}_{t-1} + \left(\frac{g}{b} - \frac{\beta^{-1}\kappa \sigma_y \rho_g}{\phi_z + \sigma_y + \phi_\pi \kappa}\right)\hat{g}_t. \end{aligned} \tag{A.3}$$

In case of the forward data rules $\hat{R}_t = \phi_\pi E_t \hat{\pi}_{t+1} + \phi_z \hat{E}_t z_{t+1}$ solve (19) and (20) for

$$\hat{\pi}_t = (\beta + \kappa \sigma_y^{-1}(1 - \phi_\pi)) E_t \hat{\pi}_{t+1} + \kappa (1 - \sigma_y^{-1}\phi_z) E_t \hat{z}_{t+1} + \kappa \rho_g \hat{g}_t. \tag{A.4}$$

Substitute $\hat{\pi}_t$ in (21) to obtain

$$\hat{b}_t \quad = \quad -(1 + \kappa \beta^{-1}\sigma_y^{-1}(1 - \phi_\pi)) E_t \hat{\pi}_{t+1} - \kappa \beta^{-1}(1 - \sigma_y^{-1}\phi_z) E_t \hat{z}_{t+1} \tag{A.5}$$

$$+ \quad (\beta^{-1} - \gamma_1)\hat{b}_{t-1} + \beta^{-1}\hat{R}_{t-1} + \left(\frac{g}{b} - \beta^{-1}\kappa \rho_g\right)\hat{g}_t. \tag{A.6}$$

In order to get a second order form in $(\hat{\pi}_t, \hat{z}_t, \hat{b}_t)$ we need to replace $\hat{R}_{t-1}$. Therefore solve (19) and (20) for

$$E_t \hat{\pi}_{t+1} \quad = \quad \beta^{-1}\left(\hat{\pi}_t - \kappa \hat{z}_t\right), \tag{A.7}$$

$$E_t \hat{z}_{t+1} \quad = \quad \frac{\phi_\pi - 1}{\beta(\sigma_y - \phi_z)} \hat{\pi}_t + \frac{\kappa(1 - \phi_\pi) + \beta \sigma_y}{\beta(\sigma_y - \phi_z)} \hat{z}_t - \frac{\sigma_y \rho_g}{\sigma_y - \phi_z} \hat{g}_t. \tag{A.8}$$

These expressions can be used to transform any forward data rule in an equivalent contemporaneous data rule $\hat{R}_t = \varphi_\pi \hat{\pi}_t + \varphi_z \hat{z}_t + \varphi_g \hat{g}_t$, where

$$\varphi_\pi = \beta^{-1}\left(\phi_\pi + \phi_z \frac{\phi_\pi - 1}{\sigma_y - \phi_z}\right), \quad \varphi_z = \beta^{-1}\left(\phi_z \frac{\kappa(1-\phi_\pi) + \beta\sigma_y}{\sigma_y - \phi_z} - \phi_\pi \kappa\right). \tag{A.9}$$

Substituting the equivalent contemporaneous data rule in $(A.6)$ yields the desired form for the bond equation. With $\hat{R}_t = \phi_\pi \hat{\pi}_{t-1} + \phi_z \hat{z}_{t-1}$ equations (19) and (20) can once again be solved for inflation

$$\hat{\pi}_t = (\beta + \kappa\sigma_y^{-1}) E_t \hat{\pi}_{t+1} + \kappa E_t \hat{z}_{t+1} - \kappa\sigma_y^{-1}(\phi_\pi \hat{\pi}_{t-1} + \phi_y \hat{z}_{t-1}) + \kappa\rho_g \hat{g}_t. \tag{A.10}$$

Combining this equation for $\hat{\pi}_t$ with (21) we can derive an equivalent expression for bonds in the lagged data case:

$$\begin{aligned} \hat{b}_t &= -(1+\kappa\beta^{-1}\sigma_y^{-1}) E_t \hat{\pi}_{t+1} - \kappa\beta^{-1} E_t \hat{z}_{t+1} + \kappa\beta^{-1}\sigma_y^{-1}(\phi_\pi \hat{\pi}_{t-1} + \phi_y \hat{z}_{t-1}) \\ &+ (\beta^{-1} - \gamma_1)\hat{b}_{t-1} + \beta^{-1}\hat{R}_{t-1} + \left(\frac{g}{b} - \beta^{-1}\kappa\rho_g\right)\hat{g}_t. \end{aligned} \tag{A.11}$$

A.1.2. Eigenvalue Criteria.

A.1.2.1. *Roots in the Unit Disk.* Let $p(z)$ be the polynomial

$$p(z) := a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0 \tag{A.12}$$

and $p^*(z)$ be the corresponding reciprocal polynomial given by

$$p^*(z) := \bar{a}_0 z^n + \bar{a}_1 z^{n-1} + \ldots + \bar{a}_n, \tag{A.13}$$

where $\bar{a}_i$ is the complex conjugate of a_i. Define the Schur transform of p by

$$T p(z) := \bar{a}_0 p(z) - a_n p^*(z) \tag{A.14}$$

and the iterated Schur transforms $T^k(p)$ recursively for all $k > 1$ by $T^k(p) := T(T^{k-1}p)$. Define $\gamma_k := T^k p(0)$. Then Henrici (1988) proofs the following two results:

THEOREM A.1.1. *Let p be a polynomial of degree n, $p \neq 0$. Then all roots lie outside the closed unit disk $|z| \leq 1$ if and only if $\gamma_k > 0, k = 1, \ldots, n$.*

PROOF. See Henrici (1988) p.493 theorem 6.8b. □

THEOREM A.1.2. *Let γ_k satisfy $\gamma_k \neq 0, k = 1, \ldots, n$. Define*

$$\pi_k := \prod_{j=1}^{k} \gamma_j. \tag{A.15}$$

Then there are r roots inside the unit circle and $n - r$ roots outside the unit circle if and only if r products are negative and $n - r$ are positive.

PROOF. The proposition follows directly from Henrici (1988) p.494 proposition 6.8c and the following discussion, which demonstrates that the number of negative elements in the sequence $\{\pi_k\}$ is equal to $\sum_{j=1}^{m}(-1)^{j-1}(n+1-k_j)$. □

Note that the proposition is not applicable if $\gamma_k = 0$ for some k. This is the case whenever p is self reciprocal, i.e. its roots are pairwise symmetric to the unit circle.[12] For a polynomial $p(\lambda) = \lambda^2 + c_1\lambda + c_2$ of degree two we have

$$\gamma_1 = c_2^2 - 1 \quad \gamma_2 = (c_2-1)^2\big((c_2+1)^2 - c_1^2\big).$$

Then the following corollary is a direct application of proposition $A.1.2$.

COROLLARY A.1.3. *If p is a polynomial of degree two given by* $p(\lambda) = \lambda^2 + c_1\lambda + c_2$ *and p is not self–reciprocal, then:*[13]

- *both roots are inside the unit circle if and only if*

$$|c_2| < 1, \tag{A.16}$$
$$|c_1| < c_2 + 1. \tag{A.17}$$

- *both roots are outside the unit circle if and only if*

$$|c_2| > 1, \tag{A.18}$$
$$|c_1| < |c_2 + 1|. \tag{A.19}$$

- *one root is inside and one root is outside the unit circle if and only if*

$$|c_1| > |c_2 + 1|. \tag{A.20}$$

Regarding polynomials of degree three $p(\lambda) = \lambda^3 + c_1\lambda^2 + c_2\lambda + c_3$ we only need the conditions for the case where one root is inside and two roots are outside the unit circle. The following proposition is taken from Woodford (2003b, p.673).

PROPOSITION A.1.4. *There is one root inside and two outside the unit circle if and only if either of the following three cases holds:* [14]
(Case I)

$$1 + c_1 + c_2 + c_3 < 0, \tag{A.21}$$
$$-1 + c_1 - c_2 + c_3 > 0. \tag{A.22}$$

[12]One special case of this symmetry occurs when some roots lie on the unit circle.

[13]The property that p is not self–reciprocal means that the roots of the polynomial cannot be grouped in pairs that lie symmetric to the unit circle (cf. Henrici, 1988, p.496). Not however that this is a non–generic case. Moreover the first statement is true even in this case. (Maybe the other two cases go through as well with a self–reciprocal polynomial in the special case of a degree two polynomial.

[14]The conditions are sufficient but only generically necessary.

(Case II) and (Case III) both require

$$\begin{aligned} 1+c_1+c_2+c_3 &> 0, & \text{(A.23)} \\ -1+c_1-c_2+c_3 &< 0 & \text{(A.24)} \end{aligned}$$

and in addition (Case II)

$$c_3^2 - c_1c_3 + c_2 - 1 > 0 \tag{A.25}$$

or (Case III) [15]

$$|c_1| > 3. \tag{A.26}$$

A.1.2.2. *Real Parts Smaller One.* E–stability requires that the real parts of the roots of the matrices (33) to (35) are smaller than one. The following result is known as the Routh–Hurwitz theorem:

THEOREM A.1.5. *Consider a real $n \times n$ matrix A with characteristic equation $\lambda^n + b_1\lambda^{n-1} + \ldots + b_{n-1}\lambda + b_n = 0$. Then the eigenvalues λ have negative real parts if and only if $\forall k = 1 \ldots n : \Delta_k > 0$, where*

$$\Delta_k = \det\left(\begin{bmatrix} b_1 & 1 & 0 & 0 & 0 & 0 & \ldots & 0 \\ b_3 & b_2 & b_1 & 1 & 0 & 0 & \ldots & 0 \\ b_5 & b_4 & b_3 & b_2 & b_1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{2k-1} & b_{2k-2} & b_{2k-3} & b_{2k-4} & b_{2k-5} & b_{2k-6} & \ldots & b_{2k-1} \end{bmatrix}\right). \tag{A.27}$$

PROOF. Gradshteyn and Ryzhik 2000, p. 1076 □

The real parts of a matrix A are less than one whenever the real parts of $A - I$ are negative and thus the proposition can be used to derive conditions for E–stability. For a (2×2) matrix A straightforward calculations show that a necessary and sufficient condition for the eigenvalues to have negative real parts is that $\det(A) > 0$ and $\mathrm{tr}(A) < 0$.

A.1.3. Determinacy.

A.1.3.1. *No Bonds–Contemporaneous Data Rules.*

PROOF OF PROPOSITION 1.4.2. The coefficients of the characteristic polynomial of $A^{-1}B$ are

$$c_1 = -\frac{\beta(\phi_z + \sigma_y) + \kappa + \sigma_y}{\phi_z + \sigma_y + \kappa\phi_\pi}, \quad c_2 = \frac{\sigma_y\beta}{\phi_z + \sigma_y + \kappa\phi_\pi}. \tag{A.28}$$

For determinacy we need both roots to be inside the unit circle, which is equivalent to the conditions $(A.16)$ and $(A.17)$. Using the assumptions $\kappa > 0$, $\sigma_y > 0$, $0 < \beta < 1$, $\phi_\pi \geq 0$ and $\phi_z \geq 0$ it is easy to see that $(A.16)$ is always fulfilled, while $(A.17)$ is true if and only if (39) holds. □

[15] Note that (Case II) and (Case III) are not disjunct. As noted by Woodford (2003b) this can be achieved by demanding that $c_3^2 - c_1c_3 + c_2 - 1 < 0$ in (Case III).

A.1.3.2. *No Bonds–Forward Data Rules.*

PROOF OF PROPOSITION 1.4.4. The coefficients of the characteristic polynomial of $A^{-1}B$ are

$$c_1 = \sigma_y^{-1}(\kappa(\phi_\pi - 1) + \phi_z) - 1 - \beta, \quad c_2 = (1 - \sigma_y^{-1}\phi_z)\beta. \tag{A.29}$$

For determinacy we need both roots to be inside the unit circle, which is equivalent to the conditions (A.16) and (A.17). It can easily be verified that these conditions are equivalent to (44), (43) and

$$\phi_z < \sigma_y(1 + \beta^{-1}). \tag{A.30}$$

Note that (43) is equivalent to $1 - \phi_\pi < \frac{1-\beta}{\kappa}\phi_z$. Rewriting (44) yields $\phi_z < 2\sigma_y + \frac{\kappa}{1+\beta}(1 - \phi_\pi)$ and thus with the previous result $\phi_z < 2\sigma_y + \frac{1-\beta}{1+\beta}\phi_z$ which is equivalent to (A.30). As a consequence determinacy is equivalent to (44) and (43). □

A.1.3.3. *No Bonds–Lagged Data Rules.*

PROOF OF PROPOSITION 1.4.6. In order to be able to apply the conditions of proposition A.1.4 look at the characteristic polynomial of $B^{-1}A$ given by $\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3$, where

$$c_1 = -\frac{\sigma_y\beta + \kappa + \sigma_y}{\sigma_y\beta}, \quad c_2 = \frac{\sigma_y - \phi_z\beta}{\sigma_y\beta}, \quad c_3 = \frac{\phi_\pi\kappa + \phi_z}{\sigma_y\beta}. \tag{A.31}$$

It can be verified easily that (Case I) is equivalent to (49) and (48). Moreover (Case II) and (Case III) are satisfied whenever the reverse inequalities (51) and (50) hold and in addition one of the inequalities (A.25) or (A.26) are satisfied respectively. Inequality (A.25) is equivalent to

$$\kappa^2\phi_\pi^2 + \kappa(\sigma_y(1+\beta) + 2\phi_z + \kappa)\phi_\pi + \phi_z^2 + (\kappa - \sigma_y(\beta^2 - \beta - 1))\phi_z + \sigma_y^2\beta(1-\beta) > 0. \tag{A.32}$$

Recall that by assumption $\phi_\pi \geq 0$, $\phi_z \geq 0$, $\kappa > 0$, $\sigma_y > 0$ and $0 < \beta < 1$. Thus condition (A.32) is always satisfied, so that (Case II) and (Case III) are equivalent to (51) and (50). □

A.1.3.4. *Positive Bonds–Contemporaneous Data Rules.*

PROOF OF PROPOSITION 1.5.3. Note from (55) that $A^{-1}B$ is block diagonal and thus the eigenvalues are given by $(\beta^{-1} - \gamma_1)^{-1}$ and the eigenvalues of the upper left (2×2) block. This block in turn is identical to (38).

Passive fiscal policy (56) is equivalent to $(\beta^{-1} - \gamma_1)^{-1} > 1$. Therefore determinacy requires both remaining roots, the eigenvalues of (38), to be inside the unit circle. By proposition 1.4.2 this is equivalent to (39).

Now consider the case of active fiscal policy (58), which is equivalent to $(\beta^{-1} - \gamma_1)^{-1} < 1$. In this case one eigenvalue of (38) has to be inside and one has to be outside the unit circle, which is equivalent to condition $(A.20)$. Condition $(A.20)$ in turn is equivalent to one of the following two conditions being true

$$(c_1 < 1 + c_2) \wedge (-c_1 > 1 + c_2), \tag{A.33}$$

$$(-c_1 < 1 + c_2) \wedge (c_1 > 1 + c_2). \tag{A.34}$$

where c_1 and c_2 are the coefficients of the characteristic polynomial of (38), given by $(A.28)$. It is easily verified that the first region of determinacy $(A.33)$ is given by (59) whereas the second region of determinacy $(A.34)$ is empty. □

A.1.3.5. *Positive Bonds–Forward Data Rules.*

PROOF OF PROPOSITION 1.5.6. Again $A^{-1}B$ (67) is block diagonal and thus the eigenvalues are $(\beta^{-1} - \gamma_1)^{-1}$ and the eigenvalues of the upper left (2×2) block. This block in turn is identical to (42).

Under passive fiscal policy (68) we have $(\beta^{-1} - \gamma_1)^{-1} > 1$ and so both eigenvalues of (42) have to be inside the unit circle. By proposition 1.4.4 this is equivalent to (44) and (43).

Active fiscal policy (71) is equivalent to $(\beta^{-1} - \gamma_1)^{-1} < 1$. Thus for determinacy one of the remaining two roots has to be inside and the other one outside the unit circle, which, as seen in the previous proof, is true if and only if conditions $(A.33)$ and $(A.34)$ hold, where c_1 and c_2 are the coefficients of the characteristic polynomial $(A.29)$. Straightforward algebra shows that $(A.33)$ is equivalent to (73) and (72), while $(A.34)$ is equivalent to (75) and (74). □

A.1.3.6. *Positive Bonds–Lagged Data Rules.*

PROOF OF PROPOSITION 1.5.9. Note that $A^{-1}B$ (83) is block triangular and thus the eigenvalues are given by $(\beta^{-1} - \gamma_1)^{-1}$ and the eigenvalues of the upper left (3×3) block. This block in turn is identical to (47).

If fiscal policy is passive (84) then $(\beta^{-1} - \gamma_1)^{-1} > 1$. Thus a determinate equilibrium requires that two eigenvalues of (47) are inside and one is outside unit circle, which by proposition 1.4.6 is equivalent to (88)-(87) or (86)-(85).

If fiscal policy is active (89) then $(\beta^{-1} - \gamma_1)^{-1} < 1$. For a determinate equilibrium one eigenvalue of (47) has to be inside and two outside the unit circle. The characteristic polynomial of $A^{-1}B$ is given by $\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3$ with:

$$c_1 = \frac{\sigma_y - \phi_z \beta}{\phi_\pi \kappa + \phi_z}, \quad c_2 = -\frac{\kappa + \sigma_y(1+\beta)}{\phi_\pi \kappa + \phi_z}, \quad c_3 = \frac{\sigma_y \beta}{\phi_\pi \kappa + \phi_z}. \tag{A.35}$$

Then one root is inside and two are outside the unit circle whenever one of the three cases defined by conditions $(A.21)$-$(A.26)$ holds. It is easily verified that (Case I) is equivalent to (91) and (90). It remains to show that (Case II) and (Case III) are empty. Consider first (Case II) and note that condition $(A.25)$ is equivalent to

$$\kappa^2\phi_\pi^2+\kappa(\sigma_y(1+\beta)+2\phi_z+\kappa)\phi_\pi+\phi_z^2+(\kappa-\sigma_y(\beta^2-\beta-1))\phi_z+\sigma_y^2\beta(1-\beta)<0. \tag{A.36}$$

It is easy to see that $(A.36)$ contradicts the assumptions $\phi_\pi \geq 0$, $\phi_z \geq 0$, $\kappa > 0$, $\sigma_y > 0$ and $0 < \beta < 1$, so that (Case II) is empty. For (Case III) note that condition $(A.26)$ can be written as

$$\left|\frac{-\phi_z\beta+\sigma_y}{\phi_\pi\kappa+\phi_z}\right|>3 \Leftrightarrow \left(\phi_z<-\frac{\sigma_y+3\phi_\pi\kappa}{3-\beta}\vee\phi_z<\frac{\sigma_y-3\phi_\pi\kappa}{\beta+3}\right). \tag{A.37}$$

However $\phi_z < -\frac{\sigma_y+3\phi_\pi\kappa}{3-\beta} < 0$ contradicts the assumption $\phi_z \geq 0$. Condition $(A.24)$ is equivalent to $\frac{\kappa}{1+\beta}(1-\phi_\pi)+2\sigma_y < \phi_z$. Together with $\phi_z < \dfrac{\sigma_y-3\phi_\pi\kappa}{\beta+3}$ (the second part of $(A.37)$) this implies $\frac{\kappa}{1+\beta}(1-\phi_\pi)+2\sigma_y < \frac{\sigma_y-3\phi_\pi\kappa}{\beta+3} \Leftrightarrow \phi_\pi < -\frac{1}{2\beta\kappa}[\sigma_y(2\beta^2+7\beta+5)+\kappa(\beta+3)] < 0$ contradicting the assumption $\phi_\pi \geq 0$. Thus (Case III) is empty too. □

A.1.4. Rational Expectations Equilibrium Representations.

PROOF OF PROPOSITION 1.5.4. Our solution method follows Evans and Honkapohja (2001) p.260. Rewrite the system (22) as

$$x_t = Jx_{t+1}+A^{-1}Cu_t+A^{-1}D\eta_{t+1}.$$

where $J = A^{-1}B$ is given by (55). The Jordan decomposition $J = Q\Lambda Q^{-1}$ is given by

$$\Lambda = \mathrm{diag}(1/\lambda_1, 1/\lambda_2, 1/\lambda_3)$$

and

$$Q^{-1} = \begin{bmatrix} -\frac{\mu_4}{\kappa\mu_1(\mu_3-\mu_4)} & -\frac{1}{\mu_1(\mu_3-\mu_4)} & 0 \\ \frac{\mu_3}{\kappa\mu_2(\mu_3-\mu_4)} & \frac{1}{\mu_2(\mu_3-\mu_4)} & 0 \\ -\frac{\mu_3\mu_1-\mu_4\mu_2}{\kappa\mu_1\mu_2(\mu_3-\mu_4)} & -\frac{\mu_1-\mu_2}{\mu_1\mu_2(\mu_3-\mu_4)} & 1 \end{bmatrix},$$

where $\mu_1 = \dfrac{\beta(\lambda_1-\lambda_3)}{-\lambda_1(\beta\phi_z+\kappa)+\kappa\phi_\pi+\phi_z}$, $\mu_2 = \dfrac{\beta(\lambda_2-\lambda_3)}{-\lambda_2(\beta\phi_z+\kappa)+\kappa\phi_\pi+\phi_z}$, $\mu_3 = \lambda_1\beta - 1$ and $\mu_4 = \lambda_2\beta - 1$. Recall that the endogenous variables $x_t = (\hat{\pi}_t, \hat{y}_t, \hat{b}_t)'$ contain two free variables and one predetermined variable. To obtain a determinate solution we need exactly two restrictions. If $|\lambda_i| > 1$ then stationarity requires that $(Q^{-1})_{i,\cdot}x_t + k_i\hat{g}_t = 0$, where $(Q^{-1})_{i,\cdot}$ is the ith row of Q^{-1}. Thus a determinate equilibrium demands that two out of the three roots are outside the unit circle. The roots can be computed as $\lambda_1 = p+\sqrt{p^2-q}$, $\lambda_2 =$

$p-\sqrt{p^2-q}$ and $\lambda_3=\beta^{-1}-\gamma_1$, where $p=\frac{\sigma_y+\kappa+\beta(\sigma_y+\phi_z)}{2\beta\sigma_y}$ and $q=\frac{\sigma_y+\phi_z+\kappa\phi_\pi}{\sigma_y\beta}$. Stationarity then requires that the corresponding two of the following three restrictions hold

$$(Q^{-1})_{1,\cdot}x_t+k_1\hat{g}_t=0 \quad\Leftrightarrow\quad (\lambda_2\beta-1)\pi_t+\kappa c_t+k_1\hat{g}_t=0, \tag{A.38}$$
$$(Q^{-1})_{2,\cdot}x_t+k_2\hat{g}_t=0 \quad\Leftrightarrow\quad (\lambda_1\beta-1)\pi_t+\kappa c_t+k_2\hat{g}_t=0, \tag{A.39}$$
$$(Q^{-1})_{3,\cdot}x_t+k_3\hat{g}_t=0 \quad\Leftrightarrow\quad b_{t-1}-\frac{\mu_3\mu_1-\mu_4\mu_2}{\kappa\mu_1\mu_2(\mu_3-\mu_4)}\pi_t-\frac{\mu_1-\mu_2}{\mu_1\mu_2(\mu_3-\mu_4)}c_t+k_3\hat{g}_t=0 \tag{A.40}$$

Consider first the case where $|\lambda_1|>1$, $|\lambda_2|>1$ and $|\lambda_3|<1$. Combining $(A.38)$, $(A.39)$ and (21) yields the monetarist solution (M) given by $\bar{c}_0=0$, $\bar{c}_1=\Gamma_2$ and $\bar{c}_2$, where $\bar{c}_2$ is the unique solution to (31). The roots of $\bar{c}_1$ are $\{0,0,\lambda_3\}$ and so passive fiscal policy (56) guarantees stationarity of the monetarist solution. Proposition 1.5.3 proves that there is a unique stationary solution under the conditions (56) and (57), which thus must be equal to the monetarist solution.

Next assume $|\lambda_1|<1$, $|\lambda_2|>1$ and $|\lambda_3|>1$. Recall that by assumption $\kappa>0$, $\sigma_y>0$, $0<\beta<1$ and $\phi_z>0$. It follows that $p>1$. If $p^2-q>0$ then $|\lambda_1|>p>1$. If $p^2-q<0$ we have $q>p^2>1$, so $|\lambda_1|=\sqrt{p^2-(p^2-q)}=\sqrt{q}>1$ contradicting the assumption $|\lambda_1|<1$.

Finally assume $|\lambda_1|>1$, $|\lambda_2|<1$ and $|\lambda_3|>1$. Combining $(A.38)$, $(A.40)$ and (21) yields the fiscalist solution (F) given by $\bar{c}_0=0$, $\bar{c}_1=(\phi_\pi v,\phi_y v,\beta\lambda_3 v)$ and $\bar{c}_2$, where $\bar{c}_2$ is the unique solution to (31) and v is

$$v=\frac{1}{(\lambda_2\beta-1)\phi_z+\kappa(\lambda_3-\phi_\pi)}\begin{bmatrix}-\kappa(\lambda_2-\lambda_3)\\(\lambda_2\beta-1)(\lambda_2-\lambda_3)\\\beta^{-1}(\lambda_2\beta-1)\phi_z+\beta^{-1}\kappa(\lambda_2-\phi_\pi)\end{bmatrix}. \tag{A.41}$$

By proposition 1.5.3 there is a determinate solution in this case. By assumption $|\lambda_3|>1$, so the monetarist solution is explosive. Therefore the fiscalist solution must be the determinate solution. □

PROOF OF PROPOSITION 1.5.7. The solutions (M), $(F1)$ and $(F2)$ can be computed following the same steps as in the proof of proposition 1.5.4. Each solution suppresses two of the three eigenvalues in $A^{-1}B$ (67). The eigenvalues of $B^{-1}A$ are $\lambda_1=-p+\sqrt{p^2-q}$, $\lambda_2=-p-\sqrt{p^2-q}$ and $\lambda_3=\beta^{-1}-\gamma_1$, where $p=\frac{\kappa(\phi_\pi-1)+\phi_z-\sigma_y(1+\beta)}{2\beta(\sigma_y-\phi_z)}$ and $q=\frac{\sigma_y}{\beta(\sigma_y-\phi_z)}$.

The monetarist solution (M) is given by $\bar{c}_0=0$, $\bar{c}_1=\Gamma_2$ and $\bar{c}_2$, where $\bar{c}_2$ is the unique solution to (31). Both fiscalist solutions $(F1)$ and $(F2)$ are of the general form $\bar{c}_0=0$,

$\bar{c}_1 = (\varphi_\pi v, \varphi_y v, \beta\lambda_3 v)$ and $\bar{c}_2$, where $\bar{c}_2$ is the unique solution to (31) and [16]

$$v = \frac{1}{(\lambda_{1,2}\beta - 1)\varphi_z + \kappa(\lambda_3 - \varphi_\pi)} \begin{bmatrix} -\kappa(\lambda_{1,2} - \lambda_3) \\ (\lambda_{1,2}\beta - 1)(\lambda_{1,2} - \lambda_3) \\ \beta^{-1}(\lambda_{1,2}\beta - 1)\varphi_z + \beta^{-1}\kappa(\lambda_{1,2} - \varphi_\pi) \end{bmatrix}. \tag{A.42}$$

The two fiscalist solutions differ only in the parameter $\lambda_{1,2}$. In case of (F1) $\lambda_{1,2} = \lambda_1$ and in case of (F2) $\lambda_{1,2} = \lambda_2$. It remains to divide the determinate region given in proposition 1.5.6 among the three candidates (M), $(F1)$ and $(F2)$.

By construction the roots of the monetarist solution (M) are $\{0, 0, \lambda_3\}$, while the roots of the two fiscalist solutions are $\{\lambda_1, 0, 0\}$ and $\{0, \lambda_2, 0\}$ in case of $(F1)$ and $(F2)$ respectively. Obviously the monetarist solution (M) is stationary provided fiscal policy is passive. Therefore the determinate solution under the conditions (68)–(69) must be equal to the monetarist solution. In case of active fiscal policy note first that the monetarist solution (M) is always explosive. Thus for active fiscal policy rules the fiscalist solutions $(F1)$ and $(F2)$ are the only candidates for a determinate solution.

Consider first solution $(F1)$. In the following we will prove that under active fiscal policy (71) the conditions (72)–(73) and $\phi_z > \sigma_y$ or (74)–(75) and $\phi_z < \sigma_y$ are sufficient for $|\lambda_2| > 1$. Then, the only remaining candidate for a determinate equilibrium is the fiscalist solution $(F1)$. As proposition 1.5.3 guarantees that these conditions are also sufficient for determinacy $(F1)$ must be the determinate solution. The first step is to observe that either $q < 0$ or $q > 1$ as $\phi_z > 0$ and $0 < \beta < 1$. Now note that under (73) and $\phi_z > \sigma_y$ or (75) and $\phi_z < \sigma_y$ we have $q - 2p + 1 = -\frac{1+\beta}{\beta} \frac{\kappa(1+\beta)^{-1}(\phi_\pi - 1) + \phi_z - 2\sigma_y}{\sigma_y - \phi_z} < 0$. The following lemma establishes that this is enough to ensure $|\lambda_2| > 1$.

LEMMA A.1.6. *Let $p, q \in \mathbb{R}$. A sufficient condition for $|\lambda_2| > 1$ is* $(q - 2p + 1 < 0) \wedge [(q < 0) \vee (q > 1)]$.

PROOF. If $p^2 < q$ then $|\lambda_2| = |p + \sqrt{p^2 - q}| = \sqrt{p^2 - (p^2 - q)} = \sqrt{q}$. As $q > p^2 > 0$ it follows that $q > 1$ and so $|\lambda_2| > 1$. If $p^2 > q$ then $q - 2p + 1 < 0 \Leftrightarrow (p-1)^2 < p^2 - q \Rightarrow -p - \sqrt{p^2 - q} < -1$ and so $|\lambda_2| > 1$. □

Next consider the solution $(F2)$. Along the same lines as above we will show that under active fiscal policy (71) the conditions (72)–(73) and $\phi_z < \sigma_y$ or (74)–(75) and $\phi_z > \sigma_y$ imply $|\lambda_1| > 1$. This will identify $(F2)$ as the only candidate for a determinate equilibrium. Again the conditions are also sufficient for determinacy $(F2)$ by proposition 1.4.2 and so $(F2)$ is the determinate equilibrium. The following lemma establishes conditions analogous to lemma $A.1.6$ that are sufficient for $|\lambda_1| > 1$.

[16] Explicit formulas for the compound parameters φ_π and φ_z are given in equation $(A.9)$.

LEMMA A.1.7. *Let $p,q \in \mathbb{R}$. A sufficient condition for $|\lambda_1| > 1$ is $(q+2p+1>0) \wedge [(q<0) \vee (q>1)]$.*

PROOF. If $p^2 < q$ then $|\lambda_1| = |-p+\sqrt{p^2-q}| = \sqrt{p^2-(p^2-q)} = \sqrt{q}$. As $q > p^2 > 0$ it follows that $q > 1$ so that $|\lambda_1| > 1$. If $p^2 > q$ then $q+2p+1 < 0 \Leftrightarrow (p+1)^2 < p^2 - q \Rightarrow -p+\sqrt{p^2-q} > 1$ and so $|\lambda_1| > 1$. □

As argued above it is easy to see that either $q < 0$ or $q > 1$. Moreover $q+2p+1 = \frac{\kappa(\phi_\pi-1)+(1-\beta)\phi_z}{\beta(\sigma_y-\phi_z)}$. It is easy to see that the conditions (72) and $\phi_z < \sigma_y$ or (74) and $\phi_z > \sigma_y$ imply $q+2p+1 < 0$ and so $|\lambda_1| > 1$ by virtue of lemma A.1.7. □

A.1.5. Learnability.

A.1.5.1. *No Bonds–Contemporaneous Data Rules.*

PROOF OF PROPOSITION 1.4.3. Recall that due to the special structure ($\bar{c}_1 = \Gamma_2 = 0$) a necessary and sufficient condition for E–stability is that the real parts of the eigenvalues of $\Gamma_1 = A^{-1}B$ are less than one. By the Routh–Hurwitz proposition $A.1.5$ this is equivalent to $|A^{-1}B - I| > 0$ and $\mathrm{tr}(A^{-1}B - I) < 0$. Straightforward algebra shows that $|A^{-1}B - I| > 0$ is equivalent to (40), whereas

$$-\mathrm{tr}(A^{-1}B-I) = \frac{[\kappa(\phi_\pi - 1) + \phi_z(1-\beta)] + \kappa\phi_\pi + \phi_z + \sigma_y(1-\beta)}{\phi_z + \sigma_y + \phi_\pi\kappa} > 0.$$

The term in square brackets is positive by condition (40). By assumption $\kappa > 0$, $\phi_\pi \geq 0$, $\phi_z \geq 0$, $\sigma_y > 0$ and $0 < \beta < 1$ and therefore the last three terms are nonnegative. □

A.1.5.2. *No Bonds–Forward Data Rules.*

PROOF OF PROPOSITION 1.4.5. As in the contemporaneous data case, a necessary and sufficient condition for E–stability is that $|A^{-1}B - I| > 0$ and $\mathrm{tr}(A^{-1}B - I) < 0$. It is easily checked that condition $|A^{-1}B - I| > 0$ is equivalent to (45). Additionally $-\mathrm{tr}(A^{-1}B - I) = [\kappa(\phi_\pi - 1) + \phi_z(1-\beta)] + \beta\phi_z + \sigma_y(1-\beta) > 0$. The term in square brackets is positive by condition (45). By assumption $\phi_z \geq 0$, $\sigma_y > 0$ and $0 < \beta < 1$ and therefore the last two terms are nonnegative. □

A.1.5.3. *Positive Bonds–Contemporaneous Data Rules (Monetarist Solution).*

PROOF OF PROPOSITION 1.5.5. First note that due to the special structure of $\bar{c}_1 = \Gamma_2$ we have $\Gamma_1\bar{c}_1 = \Gamma_1\Gamma_2 = 0$. Therefore the derivatives (29)–(30) simplify to Γ_1, $\bar{c}_1' \otimes \Gamma_1$ and $\varphi' \otimes \Gamma_1$ respectively. By assumption the eigenvalues of φ are inside the unit circle and by construction the eigenvalues of $\bar{c}_1$ are $\{0, 0, \beta^{-1} - \gamma_1\}$. The eigenvalues of the Kronecker product of two matrices are the products of the eigenvalues of each matrix. Thus the real parts of the eigenvalues of all three derivatives are smaller than one if and only if the real parts of the eigenvalues of the matrices Γ_1 and $(\beta^{-1} - \gamma_1)\Gamma_1$ are smaller than one. Γ_1 is a blocktriangular matrix with the blocks $\tilde{\Gamma}_1$ and 0 on the diagonal, where $\tilde{\Gamma}_1$ is the upper left 2×2 block. Thus the eigenvalues of Γ_1 are the eigenvalues the eigenvalues of $\tilde{\Gamma}_1$ and 0.

Under passive fiscal policy (61) we have $\beta^{-1}-\gamma_1<1$. Under active monetary policy (62) the real parts of the eigenvalues of $\tilde{\Gamma}_1$ are smaller than one (cf. proposition 1.4.3) and so the real parts of the eigenvalues of Γ_1 and $(\beta^{-1}-\gamma_1)\Gamma_1$ are smaller than one. This proves the first part of the proposition.

If fiscal policy is active (63) then $\beta^{-1}-\gamma_1>1$. Thus the condition that the real parts of the eigenvalues of $(\beta^{-1}-\gamma_1)\Gamma_1$ are smaller than one implies that the real parts of the eigenvalues of Γ_1 are smaller than one. By proposition $A.1.5$ the real parts of the eigenvalues of $(\beta^{-1}-\gamma_1)\Gamma_1$ are smaller than one if and only if $\det((\beta^{-1}-\gamma_1)\tilde{\Gamma}_1-I)>0$ and $\operatorname{tr}((\beta^{-1}-\gamma_1)\tilde{\Gamma}_1-I)<0$. Straightforward algebra shows that these two conditions can be written as (64) and (65). □

A.1.5.4. *Positive Bonds–Forward Data Rules (Monetarist Solution).*

PROOF OF PROPOSITION 1.5.8. The proof parallels the proof of proposition 1.5.5. The matrix Γ_1 is again blocktriangular. Denote the upper left 2×2 block by $\tilde{\Gamma}_1$, then the lower right block is 0. The same argument as in the proof of proposition 1.5.5 leads to the conclusion that the monetarist solution is E–stable under passive fiscal policy (77) and active monetary policy (78). Provided fiscal policy is active (79) E–stability is again equivalent to the conditions $\det((\beta^{-1}-\gamma_1)\Gamma_1-I)>0$ and $\operatorname{tr}((\beta^{-1}-\gamma_1)\Gamma_1-I)<0$, which can be written as (80) and (81). □

A.2. Appendix to Chapter 2

DEFINITION A.2.1. For $x\in\mathbb{R}^n$ and $f:R^n\to\mathbb{R}^{n\times n}$ consider the autonomous system of differential equations

$$\dot{x}=f(x).$$

A set $S\subset\mathbb{R}^n$ is said to be

(1) stable if for each ε there is a δ such that all trajectories starting in $N_\delta(S)$ never leave $N_\varepsilon(S)$,
(2) attracting if the trajectories ultimately go to S, $x\to S$, which means that $\lim_{t\to\infty}\operatorname{dist}(x(t),S)=0$, where $\operatorname{dist}(x,S)=\min_{y\in S}|x-y|$.

If S is stable and attracting then it is said to be locally asymptotically stable. If S is attracting for all initial conditions, the asymptotic stability is said to be global (cf. Kushner and Yin, 2003, chapter 4.2.2).

PROOF OF PROPOSITION 2.2.8. Vectorisation of the differential equation for d yields $\operatorname{vec}(\dot{d})=(A+I\otimes\beta(b-\bar{b}))\operatorname{vec}(d)$, where $A=DT_d(\bar{b})-I$. Note that the assumptions on the eigenvalues imply that m eigenvalues are equal to zero with a full eigenspace and the remaining eigenvalues of A have negative real parts. Therefore the Jordan decomposition of A can be written as

$$A = Q\Lambda Q^{-1} \quad \text{with} \quad \Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad Q^{-1} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix},$$

where Λ_1 contains all eigenvalues with negative real parts. Let $w = Q^{-1}\text{vec}(d)$ and $B = Q^{-1}(I \otimes \beta(b-\bar{b})Q$. Partitioning w and B correspondingly we can represent the system of differential equations for vec(d) as

$$\dot{w}_1 = [\Lambda_1 + B_1(b)]w_1, \tag{A.43}$$

$$\dot{w}_2 = B_2(b)w_2, \tag{A.44}$$

with $B_1(\bar{b}) = 0$ and $B_2(\bar{b}) = 0$. Note that a, b, c and w_1 form an independent subsystem. According to standard stability theory a necessary condition for local asymptotic stability of the solution $a \equiv \bar{a}$, $b \equiv \bar{b}$, $c \equiv \bar{c}$, $w_1 \equiv 0$ is that none of the eigenvalues of the Jacobian of the subsystem has a positive real part, i.e. that the derivatives of the T-map have no eigenvalue with a real part larger than 1. A sufficient condition is that all eigenvalues of the Jacobian have negative real parts, i.e. the eigenvalues of the derivatives of the T-map have real parts smaller than 1 apart from m eigenvalues of $DT_d(\bar{b})$ equal to one.

To complete the proof we show that local asymptotic stability of the solution $a \equiv \bar{a}$, $b \equiv \bar{b}$, $c \equiv \bar{c}$, $w_1 \equiv 0$ for the subsystem a, b, c, w_1 is equivalent to local asymptotic stability of the set $\Xi_{\bar{b}}$ for the system a, b, c and d. First we introduce the notation $\xi_v = (\xi_{v,1}', \xi_{v,2}')'$, $\xi_{v,1} = (\text{vec}(a)', \text{vec}(b)', \text{vec}(c)', w_1')'$, $\xi_{v,2} = w_2$, and $\bar{\xi}_{v,1} = (\text{vec}(\bar{a})', \text{vec}(\bar{b})', \text{vec}(\bar{c})', 0')'$. Now note that for

$$M = \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix}$$

it follows that vec$(\xi') = M\xi_v$. Assume first that $\bar{\xi}_{v,1}$ is an unstable solution of the subsystem in $\xi_{v,1}$. Then for some $\varepsilon > 0$ we can find for all δ an initial value $\xi_{v,1}(0)$ with $|\xi_{v,1}(0) - \bar{\xi}_{v,1}| < \delta$ such that for some $t > 0$ we have $|\xi_{v,1}(t) - \bar{\xi}_{v,1}| > \varepsilon|M^{-1}|$. It follows that

$$\begin{aligned} \text{dist}(\xi(t), \Xi_{\bar{b}}) &= \min_{\bar{\xi} \in \Xi_{\bar{b}}} |\text{vec}(\xi(t)') - \text{vec}(\bar{\xi}')| = \min_{\bar{\xi}_{v,2} \in \mathbb{R}^m} |M(\xi_v(t) - \bar{\xi}_v)| \\ &\geq \min_{\bar{\xi}_{v,2} \in \mathbb{R}^m} |\xi_v(t) - \bar{\xi}_v|/|M^{-1}| = |\xi_{v,1}(t) - \bar{\xi}_{v,1}|/|M^{-1}| > \varepsilon, \end{aligned}$$

so $\Xi_{\bar{b}}$ is unstable. Now assume that $\bar{\xi}_{v,1}$ is an asymptotically stable solution of the subsystem in $\xi_{v,1}$. Then there is a $\delta > 0$ for each $\varepsilon > 0$ such that for all initial values $|\xi_{v,1}(0) - \bar{\xi}_{v,1}| < \delta$ we have $|\xi_{v,1}(t) - \bar{\xi}_{v,1}| < \varepsilon/|M|$. Therefore

$$\text{dist}(\xi(t), \Xi_{\bar{b}}) = \min_{\bar{\xi}_{v,2} \in \mathbb{R}^m} |M(\xi_v(t) - \bar{\xi}_v)| \leq |M||\xi_{v,1}(t) - \bar{\xi}_{v,1}| < \varepsilon,$$

so $\Xi_{\bar{b}}$ is stable. Now it follows immediately that

$$\lim_{t\to\infty} \text{dist}(\xi(t), \Xi_{\bar{b}}) \quad \leq \quad |M| \lim_{t\to\infty} |\xi_{v,1}(t) - \bar{\xi}_{v,1}| = 0,$$

so that $\Xi_{\bar{b}}$ is also attracting. The last equality uses the fact that $\bar{\xi}_{v,1}$ is attracting for the subsystem in $\xi_{v,1}$. □

PROOF OF PROPOSITION 2.2.11. The purely forward looking version of the model (93) can be written in first order form as

$$\begin{bmatrix} y_n \\ w_n \end{bmatrix} = \begin{bmatrix} \beta & \kappa\rho^{-1} \\ 0 & \rho^{-1} \end{bmatrix} \begin{bmatrix} y_{n+1} \\ w_{n+1} \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & \rho^{-1} \end{bmatrix} \begin{bmatrix} \eta_{n+1} \\ \vartheta_{n+1} \end{bmatrix}$$

or

$$x_n = Jx_{n+1} + L\xi_{n+1}.$$

Our solution method follows closely Evans and Honkapohja (2001) p. 261. Let $J = Q\Lambda Q^{-1}$ be the Jordan normal form of J, where Λ is a block diagonal matrix. We partition $(y_n', w_n')'$ and Q^{-1} conformably as $(y_{1,n}', y_{2,n}', w_n')'$ and

$$Q^{-1} = \begin{bmatrix} Q_{11}(1,1) & Q_{11}(1,2) & Q_{12}(1) \\ Q_{11}(2,1) & Q_{11}(2,2) & Q_{12}(2) \\ Q_{21}(1) & Q_{21}(2) & Q_{22}(2) \end{bmatrix}.$$

The system becomes

$$Q^{-1} \begin{bmatrix} y_{1,n} \\ y_{2,n} \\ w_n \end{bmatrix} = \begin{bmatrix} \Lambda_1(1) & 0 & 0 \\ 0 & \Lambda_1(2) & 0 \\ 0 & 0 & \Lambda_2 \end{bmatrix} Q^{-1} \begin{bmatrix} y_{1,n+1} \\ y_{2,n+1} \\ w_{n+1} \end{bmatrix} + Q^{-1} \begin{bmatrix} L_1(1) \\ L_1(2) \\ L_2 \end{bmatrix} \xi_{n+1}.$$

The matrix $\Lambda_1(1)$ contains the unstable roots with modulus less than 1. Stationarity requires that the restrictions

$$\left[Q_{11}(1,1), Q_{11}(1,2), Q_{12}(1)\right] \begin{bmatrix} y_{1,n} \\ y_{2,n} \\ w_n \end{bmatrix} = 0 \tag{A.45}$$

hold. Due to the blockdiagonal structure of J we have $\beta = Q_{11}\Lambda_1 Q_{11}^{-1}$. Inverting Q_{11} we get the solutions

$$y_n = Q_{11}^{-1}\Big(\tilde{\Lambda}\big(Q_{11}x_{n-1} + Q_{12}w_{n-1} + L_1\xi_n\big) - Q_{12}w_n\Big), \tag{A.46}$$

where

$$\tilde{\Lambda}_1 = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_1(2)^{-1} \end{bmatrix}.$$

It follows from $(A.46)$ that $\bar{b} = Q_{11}\tilde{\Lambda}_1 Q_{11}^{-1}$. From proposition 2.2.10 we know that a necessary condition for E-stability is that all eigenvalues of $DT_b(\bar{b}) = \bar{b}' \otimes \beta + I \otimes \beta\bar{b}$ have negative real parts. We will show below that $DT_b(\bar{b})$ has at least one eigenvalue equal to 2, so that this condition is violated. Use the decompositions $\beta = Q_{11}\Lambda_1 Q_{11}^{-1}$ and $\bar{b} = Q_{11}\tilde{\Lambda}_1 Q_{11}^{-1}$ to rewrite $DT_b(\bar{b})$ as

$$\begin{aligned} DT_b(\bar{b}) &= (Q_{11}\tilde{\Lambda}_1 Q_{11}^{-1})' \otimes Q_{11}\Lambda_1 Q_{11}^{-1} + I \otimes Q_{11}\Lambda_1\tilde{\Lambda}_1 Q_{11}^{-1} \\ &= Q_{11}^{-1\prime}\tilde{\Lambda}_1 Q_{11}{}' \otimes Q_{11}\Lambda_1 Q_{11}^{-1} + I \otimes Q_{11}\Lambda_1\tilde{\Lambda}_1 Q_{11}^{-1} \\ &= (Q_{11}^{-1\prime} \otimes Q_{11})(\tilde{\Lambda}_1 \otimes \Lambda_1 + I \otimes \Lambda_1\tilde{\Lambda}_1)(Q_{11}{}' \otimes Q_{11}^{-1}) \end{aligned}$$

Noting that $Q_{11}^{-1\prime} \otimes Q_{11}$ is the inverse of $Q_{11}{}' \otimes Q_{11}^{-1}$ it follows that the spectrum of $\bar{b}' \otimes \beta + I \otimes \beta\bar{b}$ is the same as that of $\tilde{\Lambda}_1 \otimes \Lambda_1 + I \otimes \Lambda_1\tilde{\Lambda}_1$. Let k be the order of indeterminacy (which is also the dimension of $\Lambda_1(2)$), i.e. the number of free variables (or the dimension of y_n) less the number of restrictions in $(A.45)$ (or the dimension of $\Lambda_1(1)$). The lower right block of the matrix $\tilde{\Lambda}_1 \otimes \Lambda_1 + I \otimes \Lambda_1\tilde{\Lambda}_1$ is given by $\Lambda_1(2)^{-1} \otimes \Lambda_1 + I \otimes \Lambda_1\tilde{\Lambda}_1$, which in turn has the structure

$$\begin{bmatrix} \Lambda_1/\lambda_1 + \Lambda_1\tilde{\Lambda}_1 & 0 & \cdots & 0 \\ 0 & & \ddots & \vdots \\ \vdots & \ddots & & 0 \\ 0 & \cdots & 0 & \Lambda_1/\lambda_k + \Lambda_1\tilde{\Lambda}_1 \end{bmatrix} \tag{A.47}$$

where λ_i is the i-th diagonal element of $\Lambda_1(2)$. The i-th block of the previous matrix in turn has the structure

$$\begin{bmatrix} \Lambda_1(1)/\lambda_i & 0 \\ 0 & \Lambda_1(2)/\lambda_i + I \end{bmatrix} \tag{A.48}$$

So there are k blocks of the form $\Lambda_1(2)/\lambda_i + I$, $i = 1 \ldots k$, the i-th diagonal element of which is equal to 2. Thus there are k eigenvalues of $DT_b(\bar{b})$ equal to 2 so that all general form representations with $k > 0$ (i.e. $\bar{b} \neq 0$) are E-unstable. □

PROOF OF PROPOSITION 2.3.1. Proposition 2.3.1 is a special case of Delyon (1996) theorem 4. We use results from the book Benveniste et al. (1990) to show that the assumptions of Delyon (1996) theorem 4 are satisfied under our assumptions (A), (B), (Stab) and (Proj). From now on we refer to Benveniste et al. (1990) as BMP. First consider assumptions (MK1-3) of Delyon (1996). BMP show on page 291 that given assumptions (B1-2) the first two inequalities of (MK1) hold for all $q \geq 1$. Let $g(x)$ be a function on $\mathbb{R}^k$. Given $q \geq 0$ define

$$[g]_q = \sup_{x_1 \neq x_2} \frac{|g(x_1) - g(x_2)|}{(x_1 - x_2|[1 + |x_1|^q + |x_2|^q)}. \tag{A.49}$$

If $g(x) = |x|^q$ then $[g]_{q-1} < \infty$ for all $q > 0$ (cf. BMP p.290). So by lemma 9(c) p.268 of BMP the third inequality of (MK1) holds for all $q > 0$. Thus we have established that (MK1) holds for all $q > 0$. (MK2) is the same as our assumption (A2). By assumption (A3) $H(\theta, x)$ is twice continuously differentiable with bounded second derivatives on Q. Every continuously differentiable function with bounded partial derivatives satisfies a Lipschitz condition. This implies that there exist constants L_1, L_2 such that for all$\theta, \theta' \in Q$ and all $x, x' \in \mathbb{R}^l$

$$\begin{aligned} |\partial H(\theta,x)/\partial x - \partial H(\theta,x')/\partial x'| &\leq L_1|x-x'|, \\ |H(\theta,0) - H(\theta',0)| &\leq L_2|\theta-\theta'|, \\ |\partial H(\theta,x)/\partial x - \partial H(\theta',x)/\partial x| &\leq L_2|\theta-\theta'|. \end{aligned}$$

As argued by Evans and Honkapohja (2001) on p.129 in the proof of lemma 6.2 this implies that

$$\begin{aligned} |H(\theta,x) - H(\theta,x')| &\leq L_1|x-x'|(1+|x|+|x'|), \\ |H(\theta,x) - H(\theta',x)| &\leq L_2|\theta-\theta'|(1+|x|). \end{aligned}$$

It is now easy to see that $H(\theta,x)$ is of class $Li^1(Q)$ as defined by BMP on p.262. As a consequence we can apply proposition 10 p.270 of BMP which guarantees that for every θ the Markov chain (118) has a unique invariant measure μ_θ and that for some p_1, p_2

$$\begin{aligned} |v_\theta(x)| &\leq C(1+|x|^{p_1}), \\ |\Pi_\theta v_\theta(x) - \Pi_{\theta'} v_{\theta'}(x)| &\leq C|\theta-\theta'|(1+|x|^{p_2}). \end{aligned}$$

It follows that (MK1-3) hold for $p = \max\{p_1, p_2\}$ and any $q > 2p$. Together with the assumptions (A4) and (Proj) on the projections Delyon (1996) shows in the proof of theorem 4 that X_n is bounded in L_q for all $q > 0$. Next let $e_n = \gamma_n \rho_n(\theta_{n-1}, X_n)$ and note that

$$\lim_{k\to\infty} \sum_{n=0}^{k} \gamma_n e_n = \sum_{n=0}^{\infty} \gamma_n^2 |\rho_n(\theta_{n-1}, X_n)| \leq C_2 \sum_{n=0}^{\infty} \gamma_n^2 (1+|X_n|^{p_2}) < \infty \quad \text{a.s.} \tag{A.50}$$

because the expectation is finite. So e_n satisfies the A-stability condition (cf. Delyon, 1996, equation 6).

Thus assumption 1 of Delyon (1996, theorem 4) holds because (MK1-3) hold and due to our assumption (Proj). Assumption 2 corresponds to our assumption (A1). Given η and p we can choose q sufficiently large that $q/p\eta \geq 2$, because (MK1-3) hold for all $q > 0$. Thus $\sum \gamma_n^{q/p\eta} < \sum \gamma_n^2 < \infty$ by assumption (A1). From this and assumption (A4) it follows that assumption 3 of Delyon (1996, theorem 4) holds. Assumption 4 is incorporated in our assumption (Proj). Finally, $e_n = \gamma_n \rho_n(\theta_{n-1}, X_n)$ satisfies the A-stability condition (cf. Delyon, 1996, equation 6). So all assumptions of Delyon (1996, theorem 4) are satisfied and the result follows. $\square$

PROOF OF PROPOSITION 2.3.3. Let $\theta_n = \text{vec}(\xi_n, R_n)$, $\gamma_n = n^{-1}$,

$$\begin{aligned} H(\theta_n, X_n) &= \text{vec}(R_{n-1}^{-1} z_{n-1} z_{n-1}' (T(\xi_{n-1}) - \xi_{n-1}), z_n z_n' - R_{n-1}), & \text{(A.51)} \\ \rho_n(\theta_n, X_n) &= -\frac{n}{n+1} \text{vec}(0, z_n z_n' - R_{n-1}). & \text{(A.52)} \end{aligned}$$

With these definitions the learning algorithm (124) can be written in the form of the general recursive stochastic algorithm (117). It is easy to verify that assumption (A1) holds for a gain $\gamma_n = n^{-1}$. It is also easy to see that assumption (A2) and (A3) hold for the functions H and ρ_n for all $\theta \in Q$, where Q is some compact subset of $\mathbb{R}^k$. (A4) holds by assumption on the projections ψ_n. The stochastic process (126) for the state vector $X_n' = (1, y_{n-1}', w_n', \zeta_n', y_{n-2}', w_{n-1}', \zeta_{n-1}')$ can be written in the form (118) with the definitions

$$A(\theta) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ T_a(a,b) & T_b(b) & T_c(b,c) & T_d(b,d) & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varphi & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \end{bmatrix}, \quad B(\theta) \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and the vector of white noise disturbances $W_n' = (1, \vartheta_n', \varepsilon_n')$. Obviously W_n satisfies (B1). As the system matrices are continuous functions of θ it follows that $|A(\theta)| < K_1$ and $|B(\theta)| < K_3$ for some constants K_1, K_3 and all θ in a compact subset $Q \subset \mathbb{R}^k$. The Lipschitz conditions of $A(\theta)$ and $B(\theta)$ are also readily satisfied on a compact set. Moreover the eigenvalues of $A(\theta)$ are the same as those of $T_b(b)$, ρ and φ. The eigenvalues of φ are by construction of the common factor representation inside the unit circle, while those of ρ and $T_b(b)$ were assumed to be inside the unit circle. It follows that there exist constants K_2 and $0 < \omega < 1$ such that $|A(\theta)^n| < K_2 \omega^n$. So (B2) is also satisfied. Assumption (Proj) is assumed to hold. The set $\Theta_{\bar{b}}$ is assumed to be a locally asymptotically stable set of the mean ODE (120). Under this assumption the existence of a Lyapunov function as required by assumption (Stab) follows from the converse Lyapunov theorems of Wilson (1969) and Lin et al. (1996). Wilson (1969) provides a proof for the case of a locally asymptotically stable set, whereas Lin et al. (1996) give a more detailed exposition applied to global asymptotic stability. □

PROOF OF PROPOSITION 2.4.1. As the model is purely forward looking E-stability of the MSV solution implies E-stability of common factor sunspots (cf. corollary 2.2.9). The necessary and sufficient condition for E-stability of the MSV solution is condition (132) (cf. appendix for details). To prove the existence of sunspots rewrite the model in the first order form $Ax_n = BE_n x_{n+1} + \varepsilon_n$, where $x_n = (\hat{\pi}_n, \hat{y}_n)'$. The system matrices are

$$A = \begin{bmatrix} 1 & -\kappa \\ 0 & \sigma_y \end{bmatrix}, \quad B = \begin{bmatrix} \beta & 0 \\ 1-\phi_\pi & \sigma_y - \phi_y \end{bmatrix}. \tag{A.53}$$

The coefficients of the characteristic polynomial of $A^{-1}B$ are

$$c_1 = \sigma_y^{-1}(\kappa(\phi_\pi - 1) + \phi_y) - 1 - \beta, \quad c_2 = (1 - \sigma_y^{-1}\phi_y)\beta. \tag{A.54}$$

For indeterminacy of order one there must be one eigenvalue inside and one outside the unit circle, which is equivalent to the condition $|c_1| > |c_2 + 1|$ (cf. appendix A.2 for an overview of eigenvalue criteria). It is straightforward to show that this condition holds if and only if one of two sets of conditions is fulfilled. Either (132) and (133) or the same conditions with the inequalities reversed must hold. However, as argued above E-stability requires (132) to hold, so only the first set of conditions yields E-stable common factor representations. Instability of general form sunspot representations follows from proposition 2.2.11. □

Bibliography

Adam, Klaus, Albert Marcet and Juan Pablo Nicolini (2006): "Learning and Stock Market Volatility." Computing in Economics and Finance 2006 15, Society for Computational Economics.

Benveniste, Albert, Michel Metivier and Pierre Priouret (1990): *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, Berlin Heidelberg New York.

Blanchard, Olivier and Charles M. Kahn (1980): "The Solution of Linear Difference Equations Under Rational Expectations." *Econometrica*, 48:1305–1311.

Blanchard, Olivier and John Simon (2001): "The Long and Large Decline in U.S. Output Volatility." *Brookings Papers on Economic Activity*, 2001:135–164.

Branch, William A., John Carlson, George W. Evans and Bruce McGough (2004): "Monetary Policy, Endogenous Inattention, and the Volatility Trade-off." University of Oregon Economics Department Working Papers 2004-19, University of Oregon.

Branch, William A., John Carlson, George W. Evans and Bruce McGough (2006): "Adaptive Learning, Endogenous Inattention, and Changes in Monetary Policy." University of Oregon Economics Department Working Papers 2006-6, University of Oregon.

Bray, Margaret M. and Nathan E. Savin (1986): "Rational Expectations Equilibria, Learning, and Model Specification." *Econometrica*, 54:1129–1160.

Buiter, William H. (2002): "The Fiscal Theory of the Price Level: A Critique." *Economic Journal*, 112:459–480.

Bullard, James and In-Koo Cho (2005): "Escapist policy rules." *Journal of Economic Dynamics and Control*, 29:1841–1865.

Bullard, James B. and Kaushik Mitra (2002): "Learning About Monetary Policy Rules." *Journal of Monetary Economics*, 49:1105–1129.

Bullard, James B. and Kaushik Mitra (2007): "Determinacy, Learnability, and Monetary Policy Inertia." *Journal of Money, Credit and Banking*, 39:1177–1212.

Calvo, Guillermo (1983): "Staggered Prices in a Utility–Maximizing Framework." *Journal of Monetary Economics*, 12:383–398.

Cho, In-Koo, Noah Williams and Thomas J. Sargent (2002): "Escaping Nash Inflation." *Review of Economic Studies*, 69:1–40.

Christiano, Lawrence J. and Terry J. Fitzgerald (2000): "Understanding the Fiscal Theory of the Price Level." NBER Working Papers 7668, National Bureau of Economic Research, Inc.

Clarida, Richard, Jordi Galí and Mark Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Theory." *Quarterly Journal of Economics*, 115:147–180.

Delyon, Bernard (1996): "General results on the Convergence of Stochastic Algorithms." *IEEE Transactions on Automatic Control*, 41:1245–1255.

Evans, George and Seppo Honkapohja (1998): "Economic Dynamics with Learning: New Stability Results." *The Review of Economic Studies*, 65:23–44.

Evans, George and Seppo Honkapohja (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton.

Evans, George and Seppo Honkapohja (2003): "Expectations and the Stability Problem for Optimal Monetary Policies." *Review of Economic Studies*, 70:807–824.

Evans, George and Seppo Honkapohja (2004): "Policy Interaction, Learning and the Fiscal Theory of Prices." *CEPR Discussion Papers*, 3564.

Evans, George and Bruce McGough (2005): "Monetary Policy, Indeterminacy and Learning." *Journal of Economic Dynamics and Control*, 29:1809–1840.

Galí, Jordi (2002): "New Perspectives on Monetary Policy, Inflation and the Business Cycle." *CEPR Discussion Papers*, 3210.

Gordon, Robert J. (2005): "What Caused the Decline in US Business cycle volatility?" In Kent, Christopher and David Norman, editors, *The Changing Nature of Business Cycle Volatility*. Economics Group, Reserve Bank of Australia, Sydney.

Guidolin, Massimo and Allan Timmermann (2007): "Properties of equilibrium asset prices under alternative learning schemes." *Journal of Economic Dynamics and Control*, 31:161–217.

Henrici, Peter (1988): *Applied and Computational Complex Analysis*, volume 1. John Wiley and Sons.

Honkapohja, Seppo and Kaushik Mitra (2004): "Are non fundamental equilibria learnable in models of monetary policy?" *Journal of Monetary Economics*, 51:1743–1770.

Khan, Hashmat and Zhenhua Zhu (2002): "Estimates of the Sticky-Information Phillips Curve for the United States, Canada, and the United Kingdom." *Bank of Canda, Working Papers*, 02-19.

Kim, Chang-Jin and Charles R. Nelson (1999): "Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle." *The Review of Economics and Statistics*, 81:608–616.

Kushner, Harold J. and G. George Yin (2003): *Stochastic Approximation and Recursive Algorithms and Applications*. Springer-Verlag, Berlin Heidelberg New York.

Leeper, Eric (1991): "Equilibria under 'Active' and 'Passive' Monetary and Fiscal Policies." *Journal of Monetary Economics*, 27:129–147.

Lin, Yuandan, Eduardo D. Sontag and Yuan Wang (1996): "A Smooth Converse Lyapunov Theorem for Robust Stability." *SIAM Journal on Control and Optimization*, 34:124–160.

Mankiw, Gregory N. and Ricardo Reis (2002): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics*, 117:1295–1328.

Mankiw, Gregory N. and Ricardo Reis (2006a): "Pervasive Stickiness." *American Economic Review*, 96:164–169.

Mankiw, Gregory N. and Ricardo Reis (2006b): "Sticky Information in General Equilibrium."

Marcet, Albert and Thomas J. Sargent (1989a): "Convergence of Least–Squares Learning in Environments with Hidden State Variables and Private Information." *Journal of Political Economy*, 97:1306–1322.

Marcet, Albert and Thomas J. Sargent (1989b): "Convergence of Least–Squares Learning Mechanisms in Self–Referential linear stochastic models." *Journal of Economic Theory*, 48:337–368.

McCallum, Bennett T. (1983): "On Non-Uniqueness in Rational ExpectationsModels: An Attempt at Perspective." *Journal of Monetary Economics*, 11:134–168.

McCallum, Bennett T. (2001): "Indeterminacy, Bubbles and the Fiscal Theory of Price Level Determination." *Journal of Monetary Economics*, 47:19–30.

McConnell, Margaret M. and Gabriel Perez-Quiros (2000): "Output fluctuations in the United States: What has changed since the early 1980s?" *American Economic Review*, 90:1464–1476.

Meltzer, Allan (1993): "Commentary: The Role of Judgment and Discretion in the Conduct of Monetary Policy." In *Changing Capital Markets: Implications for Monetary Policy*. Federal Reserve Bank of Kansas City.

Metivier, Michel and Pierre Priouret (1984): "Applications of a Kushner and Clark Lemma to General Classes of Stochastic Algorithms." *IEEE Transactions on Information Theory*, 30:140–151.

Primiceri, Giorgio (2005): "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy." NBER Working Papers 11147, National Bureau of Economic Research, Inc.

Reis, Ricardo (2006): "Inattentive Producers." *Review of Economic Studies*, 73:793–821.

Rotemberg, Julio J. and Michael Woodford (1998): "An Optimization–Based Framework for the Evaluation of Monetary Policy." *NBER Technical Working Papers No.233*.

Rotemberg, Julio J. and Michael Woodford (1999): "Interest-Rate Rules in an Estimated Sticky-Price Model." In Taylor, John, editor, *Monetary Policy Rules*. University of Chicago Press, Chicago.

Rudebush, Glenn D. (2002): "Term Strucutre Evidence on Interest Rate Smoothing and Monetary Policy Inertia." *Journal of Monetary Economics*, 49:1161–1187.

Sargent, Thomas J. and Neil Wallace (1981): "Some Unpleasant Monetarist Arithmetic." *Federal Reserve Bank of Minneapolis, Quarterly Review*, 5:1–17.

Sargent, Thomas J. and Noah Williams (2005): "Impacts of Priors on Convergence and Escapes from Nash Inflation." *Review of Economic Dynamics*, 8:360–391.

Sims, Christopher A. (1999): "The Precarious Fiscal Foundations of the EMU." *De Economist*, 147:415–436.

Sims, Christopher A. (2001): "Solving Linear Rational Expectations Models." *Computational Economics*, 20:1–20.

Summers, Peter M. (2005): "What caused the Great Moderation? Some Cross–Country Evidence." *Economic Review*, pages 5–32.

Taylor, John B. (1996): "How Should Monetary Policy Respond to Shocks while Maintaining Long-Run Price Stability? – Conceptual Issues." *Achieving Price Stability, Federal Reserve Bank of Kansas City*, pages 181–195.

Uhlig, Harald (1999): "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easiliy." In Marimon, Ramon and Andrew Scott, editors, *Computational Methods for the Study of Dynamic Economies*. Oxford University Press, Oxford.

Wang, Pengfei and Yi Wen (2006): "Solving Linear Difference Systems with Lagged Expectations by a Method of Undetermined Coefficients." *Working Paper 2006-C003, Federal Reserve Bank of St.Louis*.

Williams, Noah (2001): "Escape Dynamics in Learning Models." *Ph. D. Thesis, University of Chicago*.

Wilson, F. Wesley (1969): "Smoothing Derivatives of Functions and Applications." *Transactions of the American Mathematical Society*, 139:413–428.

Woodford, Michael (1996): "Control of the Public Debt: A Requirement for Price Stability?" *NBER Working Paper 5684*.

Woodford, Michael (2003a): "Comment on: Multiple–Solution Indeterminacies in Monetary Policy Analysis." *Journal of Monetary Economics*, 50:1177–1188.

Woodford, Michael (2003b): *Interest and Prices, Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton and Oxford.

Yun, Tack (1996): "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles." *Journal of Monetary Economics*, 36:345–370.